**Cognitive Skills, Noncognitive Skills, and School-to-Work Transitions in Rural China***

Paul Glewwe, University of Minnesota

Qiuqiong Huang, University of Arkansas

Albert Park, HKUST, IZA, and CEPR

November 2016

**Abstract**

Economists have long recognized the important role of formal schooling and cognitive skills on labor market participation and wages. More recently, increasing attention has turned to the role of personality traits, or noncognitive skills. This study is among the first to examine how both cognitive and noncognitive skills measured in childhood predict educational attainment and early labor market outcomes in a developing country setting. Analyzing longitudinal data on rural children from one of China's poorest provinces, we find that both cognitive and noncognitive skills, measured when children are 9-12, 13-16, and 17-21 years old, are important predictors of whether they remain in school or enter the work force at age 17-21. The predictive power of specific skill variables differ between boys and girls. Conditioning on years of schooling, there is no strong evidence that skills measured in childhood predict wages in the early years of labor market participation.

**Key words**: cognitive skills; noncognitive skills; schooling; rural China

*We thank Orazio Attanasio, Paolo Falco, Corrado Giuletti, Jim Heckman, Sam Preston, and seminar and conference participants at the ASSA meetings in San Diego (2013), University of Chicago, China Agricultural University, Hong Kong University of Science and Technology, Institute of Fiscal Studies, Chinese University of Hong Kong, University of Illinois at Urbana-Champaign, Monash University, and Ohio State University, for comments on earlier drafts of this paper. Albert Park acknowledges support from the Hong Kong Research Grants Council General Research Fund [692713].

**Cognitive Skills, Noncognitive Skills, and School-to-Work Transitions in Rural China**


**Abstract**

Economists have long recognized the important role of formal schooling and cognitive skills on labor market participation and wages. More recently, increasing attention has turned to the role of personality traits, or noncognitive skills. This study is among the first to examine how both cognitive and noncognitive skills measured in childhood predict educational attainment and early labor market outcomes in a developing country setting. Analyzing longitudinal data on rural children from one of China's poorest provinces, we find that both cognitive and noncognitive skills, measured when children are 9-12, 13-16, and 17-21 years old, are important predictors of whether they remain in school or enter the work force at age 17-21. The predictive power of specific skill variables differ between boys and girls. Conditioning on years of schooling, there is no strong evidence that skills measured in childhood predict wages in the early years of labor market participation.



**Key words**: Cognitive skills; Noncognitive skills; School-to-Work Transition; schooling; rural China

**Cognitive Skills, Noncognitive Skills, and School-to-Work Transitions in Rural China**

## I. Introduction

Economists have studied the impact of education, especially years of schooling, on wages and other labor market outcomes since the seminal studies of Becker (1964) and Mincer (1974).[1] In general, schooling raises wages only if it generates skills that have returns in the labor market (see Hanushek, 2002, for developed country evidence). This is especially true in developing countries, where math, science and literacy skills vary widely across (and within) countries for children with the same years of schooling (Hanushek and Woessman, 2008; Das & Zajonc, 2010; Singh, 2016).

To date, most economists have focused on cognitive skills. To define those skills, a good starting point is the American Psychological Association's (2007) definition of cognition: "all forms of knowing and awareness, such as perceiving, conceiving, remembering, reasoning, judging, imagining and problem solving." More simply, one can define cognitive skills as the knowledge one has acquired, and one's ability to learn new knowledge. Recently, social scientists have begun to recognize the critical role of personality traits, or noncognitive skills, in determining labor productivity.[2] Noncognitive skills can be defined as patterns of thoughts, feelings and behavior that affect social interactions with others (Borghans, et al., 2008). Psychologists have developed the Big Five taxonomy of personality traits: openness, conscientiousness, extraversion, agreeableness, and neuroticism (emotional stability) (John and Srivastava, 2009). Of these, traits associated with conscientiousness and emotional stability (especially self-esteem and locus of control) appear to be highly correlated with job performance and wages (Almlund et al., 2011).

Evidence on noncognitive skills' impacts on schooling and labor market outcomes is much more limited than for cognitive skills.[3] While IQ and other cognitive measures appear to have greater influence on productivity for skill-intensive jobs or more educated workers, noncognitive skills matter across a wide spectrum of workers (Almlund et al., 2011; Segal, forthcoming), especially those at the bottom of the wage distribution (Lindqvist and Vestman, 2011). Lindqvist and Vestman's (2011) review of 13 studies finds that a one standard deviation improvement in measured noncognitive skills increases wages by 4 to 8%. Heckman, Stixrud, and Urzua (2006) estimate an 11.2% return to a one standard deviation increase in latent noncognitive skills in the US, and find that those skills predict adult wages about as well as cognitive skills. However, Heckman et al. (2011) analyze the same data and find that noncognitive skills strongly predict educational attainment but have little effect on wages after controlling for years of schooling. In studies of the US, the most common noncognitive skill variables used in empirical studies are psychological measures of locus of control (Rotter) and self-esteem (Rosenberg).

A major gap in the literature on noncognitive skills' impacts on employment outcomes is that nearly all studies are on developed countries, especially the US. The lack of developing country studies reflects the fact that, unlike the US National Longitudinal Study of Youth (NLSY), no developing country household survey has measured noncognitive skills during childhood and followed those children into adulthood. Panel data are needed because relating current

[1] See Card (1999 and 2001) for reviews of studies that attempt to estimate the causal impact of schooling on wages, and Glewwe (2002) for evidence on the impact of schooling on labor productivity in developing countries.

[2] Almlund et al (2011) and Lindqvist and Vestman (2011) review this literature.

[3] See Goldsmith et al. Darity (1997) for US evidence and Bowles, Gintis, and Osborne (2001) for a literature review.

noncognitive skills to employment outcomes is plagued by bias from reverse causality due to the fact that work experience may greatly influence noncognitive skills.

This paper analyzes longitudinal data on rural children in one of China's poorest provinces--Gansu. It is among the first studies to examine how both cognitive and noncognitive skills predict educational attainment and early labor market outcomes in a developing country setting using panel data spanning the school-to-work transition. Given that households in poor, rural settings often face large risks, and most jobs have low skill intensity, one might expect the absolute and relative importance of noncognitive skills to be greater in developing countries. Yet high cognitive skills may also be relatively scarce in such settings. Thus, empirical evidence on the absolute and relative importance of noncognitive skills in developing countries is of great interest.

Our results indicate that as one would expect, children with higher cognitive skills at age 9-12 are more likely to still be in school at age 17-21, and ultimately complete more years of schooling. More interestingly, even after controlling for cognitive skills, some of the children's noncognitive skills measured in this study at ages 9-12 and 13-16 are also important predictors of school-to-work transitions at age 17-21. In contrast, there is little evidence that cognitive and noncognitive skills predict wages after conditioning on years of schooling. This does not necessarily imply that those skills do not affect labor productivity; it may be that years of schooling is closely correlated with such skills and so adding noisy measures of cognitive and noncognitive skills provides little new information, or that skills have weaker impacts on wages just after entering the labor market among those with less schooling. There is suggestive evidence that contemporaneously measured noncognitive skills influence wages.

The remainder of this paper is organized as follows. The next section describes the data used, after which Section III describes the estimation strategy. The results are presented in Section IV, and a final section draws conclusions and provides suggestions for future research.

## II. Data

This study analyzes data from the Gansu Survey of Children and Families (GSCF), a longitudinal study of 2,000 children in rural villages in Gansu Province who were 9-12 years old in the year 2000. Gansu, in northwest China, is one of China's poorest provinces, consistently ranked last or second to last in rural income per capita among Chinese provinces. In 2000, its population was 25.6 million, 76% of whom lived in rural areas. Its socioeconomic and educational profiles resemble those of China's other interior provinces. Relative to China as a whole, Gansu has high illiteracy and low educational expenditures. Rural residents work primarily in subsistence farming, animal husbandry, and migrant wage labor.

This study uses data from the first three waves of the GSCF, conducted in 2000, 2004, and 2007-09. Of the children surveyed in 2000, only 9 children had never enrolled in school, and another 19 children had left school before wave 1 (June 2000). In each wave, the GSCF collected detailed data using separate questionnaires for the sampled children, their parents, their teachers, school principals, and community leaders.

The GSCF has low sample attrition. Of the original 2000 children, all but one have complete information in the first survey wave, including a variety of tests and questions that measure both cognitive and noncognitive skills. Of the 1,999 children with complete first wave information, 1,872 (93.6%) were re-interviewed in wave 2 (2004), when they were 13-16 years old.[4] All 1,872 answered questions designed to measure noncognitive skills (i.e. completed the child questionnaire) and 1,773 completed Chinese and math achievement tests.[5]

Wave 3 of the GSCF consisted of two distinct data collection efforts. First, household questionnaires were completed in mid-2007 by the children's parents; they collected information on the children's education and employment. The children were interviewed 18 months later during spring festival in late January 2009 when many who had migrated returned home to visit their families. At that time, the (former) children were 17-21 years old. If they were not at home, their parents were asked the questions in the child questionnaire on their child's education and employment, but skill measures were not collected. This study uses data from the wave 3 child questionnaire which was completed for 1,863 of the original 2,000 children. This includes 86 children for whom no data were collected in wave 2, so the 137 children with no data in wave 3 consist of 42 who were missing in wave 2 and 95 who were present in wave 2. Of these 1,863 youths, 427 did not answer the child questionnaire themselves because they had migrated and did not return home for the 2009 spring festival (240), were away for military service or higher education (83), or were away for unknown reasons (104). Of the 1,436 who were available, almost all were tested. More specifically, 1,409 (75.6% of the 1,863 with child questionnaire data) completed the wave 3 literacy test, 1,423 (76.4%) completed wave 3 tests for noncognitive skills, and 1,392 (74.7%) completed both.

**A. Data on Education, Cognitive Skills and Noncognitive Skills.** The GSCF collected detailed data on each child's educational outcomes. The household questionnaires obtained data on whether the child attended kindergarten, the age when he or she first enrolled in primary school, whether he or she repeated or skipped a grade (and, if so, which ones), highest grade completed, current enrollment, and distances to the nearest primary, middle, and high schools.

In waves 1 and 2, data were also obtained from the child's homeroom teacher, including that teacher's assessment of the child's behavior and study habits, and his or her past grades in Chinese and math. The children's mothers also were asked about the child's behavior and attitudes toward schooling. Finally, a principal questionnaire provided information about the school, and a village leader questionnaire obtained community information.

A child questionnaire was completed in each survey wave. In waves 1 and 2, in addition to questions to assess noncognitive skills (see below), children were asked about their time spent on homework and their attitudes toward education. In wave 3, as they were 17-21 years old and most had finished school, questions focused on school-to-work transitions, including work and migration histories, and the earnings from, and the characteristics of, their current jobs.

The GSCF collected cognitive and noncognitive skill data in each survey wave, although the skills covered vary by wave. They are summarized in Table 1, and described in the next paragraphs.

A general cognitive ability test was administered in wave 1. Developed by experts at the Institute for Psychology of the Chinese Academy of Social Sciences, it is comprised of four sets of questions. First is a set of miscellaneous questions of common knowledge, such as "What does the stomach do?" The second set contains abstract questions that ask what two objects have

[4] In addition to the 1,872 children re-interviewed in wave 2, some information was collected from parents for another 52 children via the household questionnaire, although the children were not re-interviewed. Since we need information from the child questionnaire, these 52 are excluded from our analysis.

[5] Students who did not complete achievement tests tended to be those who had left school and held jobs that prevented them from taking the exams that were organized in local schools during day time.

in common; an example of such a pair is “elbow” and “knee”. The third is a set of simple arithmetic problems that are read out to the children, and the fourth consists of written arithmetic questions to be answered within 30 to 75 seconds (example: “A child has 12 children’s books, he/she gives 5 of them to a friend. How many does he/she have left?”).

The Chinese and math achievement tests in waves 1 and 2 were designed by experts at the Gansu Educational Bureau to cover the official primary school curriculum. They varied by grade levels,[6] and were administered in school classrooms for currently enrolled children, and in village committee offices for out of school children. In wave 1, half the children were randomly assigned to take the Chinese test and the other half took the math test. In wave 2 all students took both tests. Chinese and math tests were not administered in wave 3; less than half of the sample was still in school in 2009, so there was no common curriculum on which to base tests.

A literacy (“life skills”) test was administered in waves 2 and 3. It was designed by an expert from the China Educational Research Institute in Beijing, modeled after the International Adult Literacy Surveys (OECD and Statistics Canada, 2000). It assesses three literacy domains: prose literacy, document literacy, and numeracy. Prose literacy consists of knowledge and skills needed to understand and use information from texts containing extended prose organized in a paragraph structure typical of editorials, news stories, brochures, pamphlets, manuals, and fiction. Document literacy focuses on knowledge and skills for locating and using information found in printed materials that contain more abbreviated language and use a variety of devices to convey meaning, such as tables, charts, graphs, indices, diagrams, maps and schematics. Numeracy is the ability to interpret, apply, and communicate mathematical information in commonly encountered situations. Unlike the wave 1 and 2 Chinese and math achievement tests, the literacy test focuses on applying literacy and numeracy skills to function effectively in society. The same test was taken by everyone, regardless of grade level. The wave 2 and 3 tests are not identical. To reflect greater education and experience at older ages, the wave 3 test put more weight on reading comprehension and less on understanding simple figures.

Measurements of noncognitive skills are constructed from sets of questions in each wave of the GSCF child questionnaire. Measures of internalizing and externalizing behavior are identical in waves 1 and 2. Internalizing behavior problems are *intra*personal in nature, such as anxiety, depression and withdrawal. Externalizing problems are *inter*personal in nature and characterized by destructive behavior, impulsivity, aggression and hyper-activity (Achenbach and Edelbrock, 1978). Child psychology research suggests that environments that destabilize a child’s sense of self control over his or her life can increase internalizing problems (Dearing et al 2006; Chorpita and Barlow, 1998), while environments that impede a child’s self-regulatory efforts, or the presence of anti-social role models, can increase externalizing problems (Evans, 2004). Among the Big Five personality traits, both internalizing and externalizing behavior measure dimensions of emotional instability (or neuroticism), as do all of our other noncognitive skill variables (resilience, depressive symptoms, self-esteem) except for educational aspirations.

To measure internalizing and externalizing behavior, children were read 36 statements and asked whether they fully agreed, agreed, disagreed, or totally disagreed with the statement. An example of a statement used for the internalizing index is: “I am shy.” An example for the

[6] In 2000, two tests were given, one for grades 3 and below and one for grades 4 and above; in 2004 separate tests were given in grades 1-2, 3-4, 5-6, and in each year of middle school.

externalizing index is: "I often lose my temper with others."[7] From the two indices, the internalizing and externalizing Item Response Theory (IRT) scores were calculated by fitting the rating scale model (RSM; Andrich 1978). They were then standardized to measure standard deviations from the mean.[8] For both, higher scores indicate more behavioral problems. All noncognitive skill measures *except* educational aspirations were similarly transformed into standardized IRT scores.

Wave 1 includes three other noncognitive skill measures. Educational aspirations are measured by whether the highest schooling level the child wants to achieve is college or above. Limited scales were also constructed for self-esteem and depressive symptoms using questions similar to, but with fewer items than, those found in the Rosenberg self-esteem and CES-D depression scales used in wave 3 (described below). The notes to Table 1 provide further details.

In wave 2, internalizing and externalizing behavior, and educational aspirations, were measured again. Also, resilience was measured using a set of questions, in the form of statements for which respondents were asked whether they agree. The statements are based on Song's (2001, 2003) adaption of Noam and Goldstein's (1998) "Resilience Inventory". Resilience is the capacity to achieve favorable outcomes despite challenges and risk factors. The Resilience Inventory (RI) was developed to be culturally sensitive and has been administered to children and adolescents in China, Israel, South Korea, Switzerland, Thailand, and the US. The GSCF adapted Song's RI for South Korea, which includes six subscales covering optimism, self-efficacy, relationships with adults, peer relationships, interpersonal sensitivity and emotional control. This was the first large-scale survey of child and adolescent resilience in rural China.

Wave 3 collected no data on internalizing or externalizing behavior, nor on resilience; instead it collected data on two other noncognitive skills: self-esteem (Rosenberg Self-Esteem Scale) and depressive symptoms (Center for Epidemiological Studies Depression Scale, CES-D). This was done because the internalizing, externalizing and resilience scales are designed for children and adolescents, while the self-esteem and depressive symptoms scales are more relevant for adults. The Rosenberg scale measures perceptions of self-worth, that is, an person's degree of approval or disapproval of himself (Rosenberg, 1965). It is widely used, and has accumulated evidence of validity and reliability. Song et al. (2011) shows that Rosenberg self-esteem scale is suitable for use in Chinese population and has been applied in numerous studies in China (e.g., Cai et al., 2009). It contains 10 statements such as "Overall, I feel satisfied with myself"; respondents are asked whether they strongly agree, agree, disagree, or strongly disagree. CES-D is often used in surveys to detect depressive symptoms. Developed at the US National Institute of Mental Health in the 1970s (Radloff, 1977), it consists of 12 statements such as "I felt that everything I did was an effort". For the last week, respondents are asked to express, on a four point scale (never, once in a while, sometimes, and frequently), the frequency of various feelings. CES-D has been validated for some Chinese population (Boey, 1999) and used in Chinese studies (Yen et al., 2000).

To aggregate multiple cognitive skills and multiple noncognitive skill measures into indices of cognitive skills (or *cognitive factors*) and noncognitive skills (or *noncognitive factors*), we employ exploratory factor analysis for each skill category in each year and include all retained

[7] Both internalizing and externality scale have high internal consistency, with a Cronbach alpha of 0.82 and 0.89 respectively. Both internal and external scales have also been implemented and validated in studies in more than 30 societies (Ivanova et al., 2007a and 2007b) as well as in other studies in China (e.g., Chen et al., 2002).

[8] The Rasch model (Rasch 1960, 1961) is a commonly used IRT model for binary responses. The partial-credit model (PCM; Masters 1982) extends the Rasch model to ordered response categories 0, 1,…, *m*. The rating scale model (RSM; Andrich 1978) is a special case of PCM. It is appropriate if the *m* response categories have the same meaning for all items and if the differences in step difficulties for different categories are the same for all items.

factors that explain more than 3% of explained variance. This method identifies latent factors that underly different skill measurements. This reduces the number of cognitive factors to 2 in both 2000 and 2004 and noncognitive factors to 3 in 2000 and 2 in 2004. For each factor in each year, the analysis produces factor loadings (or weights) for each of the individual skill measures. To provide intuition for which specific skills are most captured by each aggregate factor, in discussing each factor we note in parentheses the specific skill(s) that has(have) the largest weights, or score coefficients.[9]

**B. Employment and Wage Data.** Of the 2,000 children sampled in the year 2000, data were collected for 1,863 of them using the wave 3 child questionnaire. Of these, 847 (45.5%) were still in school, 849 (45.6%) were working, and 167 (9.0%) were neither working nor in school. Wages are not observed for those still in school, who will attain more years of education and so receive higher wages than early labor market entrants, thus observed earnings will disproportionately reflect the lower end of the education and wage distributions. Of the 849 young people working in 2009, 769 (90.6%) worked for wages and the rest were self-employed.[10] Also of interest is that 518 (61.0%) of these 849 were working in another province, and 167 (19.7%) were working in Gansu but in a different county than the one in which they had grown up.

The dependent variable in the earnings regressions presented below is the log of hourly earnings. It is computed based on the answers to three questions: monthly income from current job (including bonuses and subsidies), days worked per month, and hours worked per day.

## III. Estimation Strategy

The goal of this paper is to examine whether *noncognitive* skills predict educational outcomes, employment status and wages after controlling for age, experience, schooling and *cognitive* skills. The basic approach is to estimate standard models of these education and labor market outcomes, and add *noncognitive* skills as additional predictors. In contrast to structural approaches that aggregate many measurements of personality traits into one latent noncognitive skill, we allow for the possibility that different personality traits are distinct skills that may influence labor market outcomes differently. We take the same approach for cognitive skills. For purposes of comparison, we also estimate specifications in which multiple skill measurements are aggregated into indices of cognitive skills and noncognitive skills using exploratory factor analysis.

[9] To ease interpretation, when using the factor analysis all variables are recoded so that higher values correspond to better outcomes. We report the scoring coefficients for regressions of varimax rotated factors. The scoring coefficients of the 2000 first cognitive factor (cognition) are: cognition 0.522, Chinese 0.402, math 0.030. Scoring coefficients for the 2000 second cognitive factor (math) are: math 0.447, cognition 0.251, and Chinese -0.257. Scoring coefficients for the 2004 first cognitive factor (Chinese/math) are: Chinese 0.431, math 0.428, and literacy 0.147. Scoring coefficients for the 2004 second cognitive factor (literacy) are: literacy 0.184, Chinese 0.060, and math 0.039. Scoring coefficients for the 2000 first noncognitive factor (internalizing) are: internalizing behavior 0.536, externalizing behavior 0.393, depression 0.101, Rosenberg 0.004, and aspirations 0.002. Scoring coefficients for the 2000 second noncognitive factor (externalizing) are: externalizing 0.520, Rosenberg 0.329, aspirations 0.107, internalizing -0.199, and depression -0.427. Scoring coefficients for the 2000 third noncognitive factor (depression) are: depression 0.447, aspirations 0.307, externalizing 0.095, Rosenberg 0.081, and internalizing -0.432. Scoring coefficients for the 2004 first noncognitive factor (externalizing) are: externalizing 0.701, internalizing 0.336, aspirations -0.041, resilience -0.069. Scoring coefficients for the 2004 second noncognitive factor (resilience) are: resilience 0.490, externalizing 0.327, aspirations 0.187, and internalizing -0.436. The scoring coefficients for the 2009 noncognitive factor are: Rosenberg 0.405, and depression 0.405.

[10] In two cases, the nature of employment relationship was not reported.

To model how skills affect labor market outcomes, and to consider issues that arise when estimating those relationships, assume that different dimensions of human capital, including cognitive and noncognitive skills, affect labor productivity, and thus wages, as follows:

$$lnW_{ivt} = \propto_C \theta_{ivt}^{C} + \propto_N \theta_{ivt}^{N} + u_{vt} + \omega_{ivt} \tag{1}$$

Log wages of person $i$ in village $v$ in year $t$ is a function of cognitive ($\theta_{ivt}^{C}$) and noncognitive ($\theta_{ivt}^{N}$) skills, village characteristics such as labor market conditions ($u_{vt}$), and a residual ($\omega_{ivt}$) that reflects unobserved individual and job attributes, unmeasured skills, and measurement errors in $lnW_{ivt}$. While $\theta_{ivt}^{C}$ and $\theta_{ivt}^{N}$ can be vectors, for expositional simplicity we write them as scalars.

Equation (1) can also be used to model levels of schooling (Heckman, Stixrud, and Urzua, 2006). The outcome variable (left hand side variable) changes to $I_{ivt,s}$, which represents the benefit of schooling level $s$. An individual chooses the level of schooling that generates the highest benefit.

In a cross-sectional regression, the coefficients on the skill variables are biased if they are correlated with the village- or individual-specific errors. This can arise due to simultaneity bias, for example if employment outcomes affect cognitive and noncognitive skills. It can also reflect omitted variable bias, for example if "innate ability" affects both skills and wages or if parental preferences for child well-being affect skill formation and also influence wages because parents help their children find better jobs. Omitted variables can also include village characteristics; for example better off villages may have both high quality schools and higher wages.

A particularly serious econometric problem is errors in measured skills. Define $S_{ivt}^{k}$ to be observed skill measurements for skill type $k$ (= $C$ or $N$). Then we can write:

$$S_{ivt}^{k} = \theta_{ivt}^{k} + m_{ivt}^{k} \tag{2}$$

where $m_{ivt}^{k}$ is random measurement error. Substituting this into equation (1) yields:

$$lnW_{ivt} = \propto_C (S_{ivt}^{C} - m_{ivt}^{C}) + \propto_N (S_{ivt}^{N} - m_{ivt}^{N}) + u_{vt} + \omega_{ivt} \tag{3}$$

Random measurement errors lead to bias toward zero for coefficients of mismeasured variables.

To deal with simultaneity bias, one approach is to examine how skills measured in childhood, before individuals have entered the labor market, influences current labor market outcomes. To clarify how such lagged measurements can affect estimation of equation (1), we follow the work on the technology of skill formation by Cunha and Heckman (2008). We posit that all skills of individual $i$ in one period are a function of cognitive and noncognitive skills in the previous period ($\theta_{ivt-1}^{C}$ and $\theta_{ivt-1}^{N}$) plus household and school attributes ($X_{ivt-1}$) that influence child learning, both directly and indirectly (i.e., through their influence on specific investments parents make that can affect skill attainment):

$$\theta_{ivt}^{k} = f(\theta_{ivt-1}^{C}, \theta_{ivt-1}^{N}, X_{ivt-1}, \varepsilon_{ivt}^{k}). \tag{4}$$

where $\varepsilon_{ivt}^{k}$ is unobserved factors or idiosyncratic shocks that affect a child's skill attainment.[11]

Following Cunha and Heckman (2008), assume a linear relationship for equation (4):[12]

$$\theta_{ivt}^{k} = \beta_C^k \theta_{ivt-1}^{C} + \beta_N^k \theta_{ivt-1}^{N} + \beta_X^k X_{ivt-1} + \varepsilon_{ivt}^{k} \tag{5}$$

[11] While this specification is quite general, it still imposes strong assumptions about the dynamics of skill formation, namely that lagged values measured in period $t$-1 capture all relevant influences from earlier periods. If this is not the case, equation (4) is more accurately described as a conditional expectation.

[12] Cunha et al (2010) relax the linearity assumption using a nonparametric estimator. Our approach differs somewhat from that of Cunha and Heckman (2008) and Cunha et al (2010). First, instead of aggregating several skill measures into latent cognitive and noncognitive indices, the influence of which is estimated structurally, most of our results allow each skill measurement to independently influence labor outcomes. Second, we assume linear relationships to make identification more transparent, especially when adding village fixed effects and using instrumental variables.

It is straightforward to substitute (5) and (2) into (1) and regroup terms to arrive at the following equation for the relationship between current wages and lagged skill measurements:

$$lnW_{ivt} = (\propto_C \beta_C^C + \propto_N \beta_C^N)S_{ivt-1}^C + (\propto_C \beta_N^C + \propto_N \beta_N^N)S_{ivt-1}^N + (\propto_C \beta_X^C + \propto_N \beta_X^N)X_{ivt-1} + \propto_C \varepsilon_{ivt}^C + \propto_N \varepsilon_{ivt}^N - (\propto_C \beta_C^C + \propto_N \beta_C^N)m_{ivt-1}^C - (\propto_C \beta_N^C + \propto_N \beta_N^N)m_{ivt-1}^N + u_{vt} + \omega_{ivt} \quad (6)$$

Equation (6) has several notable characteristics. First, all skill measurements are from previous time periods, so simultaneity bias due to the impact of wages (and the error term, $u_{vt} + \omega_{ivt}$) on past skills is unlikely. Second, lagged cognitive skills affect current wages through their impact on current cognitive skills *and* current noncognitive skills, and similarly for lagged noncognitive skills. These potentially complex interactions between cognitive and noncognitive skill development mean that the relative importance of *lagged* cognitive versus noncognitive skills on current wages may not indicate the relative importance of *current* cognitive versus noncognitive skills on wages. Third, attenuation bias may arise due to measurement error in lagged skills. Fourth, as before there may still be omitted variable bias if lagged skills are correlated with unobserved individual, family, or village attributes ($u_{vt} + \omega_{ivt}$) that affect labor productivity or with unobserved factors that influence formation of current skills ($\varepsilon_{ivt}^C$ and $\varepsilon_{ivt}^N$) such as through serial correlation in the unobserved factors.

Equation (4) suggests one way to estimate equation (3): use lagged skill measurements as instruments for current skill measurements. The first-stage relationship for this method is:

$$S_{ivt}^k = \beta_C^k S_{ivt-1}^C + \beta_N^k S_{ivt-1}^N + \beta_X^k X_{ivt-1} - \beta_C^k m_{ivt-1}^C - \beta_N^k m_{ivt-1}^N + \varepsilon_{ivt}^k + m_{ivt}^k \quad (7)$$

This implies that (6) is the reduced form relationship between current wages and the instruments.

More generally, we address the econometric problems described above in several ways. First, our estimates of (6) are essentially reduced form estimates, which minimize simultaneity bias problems. Second, to estimate (3) we use past measures of skill variables as instruments for current measures to minimize simultaneity bias.[13] Third, we include village fixed effects in estimates of both (3) and (6) to avoid bias from omitted community-level factors (including omitted school quality). Fourth, to address omitted variable bias in estimates of both (3) and (6) we include a rich set of individual and family variables available in the GSCF data to control directly for likely confounding factors. Fifth, to address measurement error bias in estimating (6), we check the robustness of the results to using teacher and mother assessments to instrument for skill measurements based on child self-reports. Finally, in estimating (3) and (6) we also address measurement error and possible multicollinearity problems by replacing multiple cognitive and noncognitive skill measures with indices for each type of skill based on factor analysis.

As noted above, a human capital model would predict that only the skills learned in school or via work experience, and not years of schooling or years of work experience *per se*, should determine work productivity. However, years of schooling or work experience can have additional explanatory power if they reflect aspects of cognitive or noncognitive skills that affect worker productivity but are not captured by the skill measurements in our data. Years of schooling can also be included by appealing to a signaling model in which employers use years of schooling to infer the skills acquired by workers, especially in their first years of employment.

When wage is the outcome variable, an additional estimation issue is selection bias. Only 45% of the individuals in the sample were working for wages when interviewed in early 2009. They are unlikely to be a random sample of the original 2000 children, and our goal is to estimate a relationship that applies to the entire sample. We control for selection bias using a nonparametric sample selection method using selection propensity scores (Vella, 1998; Das,

[13] We also tried height-for-age (from wave 2) and mothers' reports of birthweight; neither had explanatory power.

Newey, and Vella, 2003). The first step is to estimate a probit model of the decision to work for wages in early 2009, and the second is the wage equation, in which we control for a function of fitted probabilities of selection.[14] In the wage regressions using 2009 skill measurements, there is an additional selection problem because we could not administer tests to migrants who did not come home for the 2009 spring festival even though their parents did provide wage information. To address this we estimate a second probit model of whether skill measurements are observed and control nonparametrically for the selection bias in the same way.

We use two variables to identify the selection correction terms (and so exclude them from the wage regression): whether the child passed the upper secondary school entrance exam, and years of bad harvests experienced by the child's family from 2000 to 2006. Children whose entrance exam scores are below a fixed threshold cannot enroll in upper secondary school (though they may be able to enroll in vocational school), which will affect their future labor supply and job options. But that event should have no predictive power on wages after controlling for their math and Chinese test scores, which are based on the official curriculum, and for other cognitive and noncognitive skills. This strategy exploits the discontinuity that occurs at the exam cutoff threshold while controlling for the effect of variation in the continuous test score variables. However, it may be that the 2000, 2004 and 2009 skill variables are noisy, which would lead to underestimation of their predictive power for wages. In addition, the 2000 and 2004 skill measurements may be outdated, so that a selection correction term generated by using whether the student passed the upper secondary entrance exam as the identifying variable could include more recent information on cognitive skills and so be correlated with the error term in the wage regression. Without correction, selection bias is expected to lead to downward bias in the coefficients on skill variables if higher unobserved ability or skill is associated with lower likelihood of working as well as higher wages. This is because those with high observed skills who work are likely to have negative unobservables and thus lower wages, flattening the wage-skill gradient. If the variable used to identify the selection correction term (i.e., passing the high school entrance exam) does contain information on skills that is not captured by the skills variables in the wage equation, applying Heckman's method may not remove all of the downward bias due to sample selection.

The second identifying variable, years of bad harvest shocks, should increase the probability of wage work if it reduces family income available to pay for education or if it increases the family's need to earn more income.[15] We posit that lagged shocks are uncorrelated with current wages after controlling for current family wealth and village fixed effects. If negative shocks adversely affected unobserved productivity even with a full set of wealth and other controls, bias on the coefficients on skill variables still arises only if the shocks are systematically correlated with the skill measurements, which seems unlikely.

Finally, briefly consider estimation of the determinants of labor force participation and years of schooling, both of which are of interest. All of the econometric issues discussed above also apply to the estimation of these relationships, with two exceptions: sample selection is not an issue; and neither years of schooling nor work experience is included as a regressor. The same methods used to address the other problems in the wage equations (village fixed effects, a rich

[14] In the reported results, we control only for the propensity score only. All results are nearly identical when we control for higher order polynomials of the propensity score.

[15] In the 2004 village leader questionnaire, leaders were asked for reasons why children dropped out of lower secondary school; they could chose multiple reasons, and 51% cited that tuition and fees were too high, which was the second most cited reason.

set of individual and family characteristics, IV estimation using lagged skill variables) are also used to estimate equations that model the determinants of labor force participation and years of schooling.

## IV. Results

In this section, we first describe the correlations among the different cognitive and noncognitive skill measures to gain insight into the information they contain. Next, we describe how the individual skill measures are combined to form index measures for cognitive and noncognitive factors. We then use the estimation results to examine the extent to which cognitive and noncognitive skills predict outcomes related to the school-to-work transition, including completed years of schooling. Finally, we assess whether cognitive and noncognitive skills predict wages.

**A. Correlations among Skills and Construction of Cognitive and Noncognitive Factors.** Evidence that cognitive and noncognitive skill variables capture different attributes is given in Table 2. In 2000, the cognitive ability test is positively correlated with the Chinese and math tests (correlation coefficients of 0.322 and 0.247, respectively). The correlation between the Chinese and math tests in 2000 cannot be calculated since children took either one or the other. In contrast, the correlations between the Chinese and math tests and the five indicators of noncognitive skills in 2000 are lower in absolute value, ranging from -0.147 to 0.120. The cognitive ability test is somewhat more correlated (in absolute value) with noncognitive skills, ranging from -0.335 to 0.214. The strongest correlations in 2000 are the among noncognitive skill measures, such as between internalizing and externalizing behavior (0.816) and between those two indicators and the depressive symptoms scale (0.628 and 0.675).[16]

Similar patterns are found for 2004 and 2009. In 2004 all children took both the Chinese and the math tests, and their correlation is quite high: 0.485. They are also highly correlated with the literacy test (0.313 and 0.310), but these three cognitive skill tests all have lower correlations with measured noncognitive skills (ranging from -0.114 to 0.244). Finally, the 2009 correlations include years of schooling, which is used in the wage regressions; it is highly correlated with the literacy test (0.493), but both years of schooling and the literacy test are far less correlated with the self-esteem scale (0.180 and 0.171) and the depressive symptoms index (-0.055 and -0.045).[17]

Correlations of both cognitive and noncognitive skills over time, shown in Table 3, are also informative. Test scores often show low correlation over time due to both measurement error and low learning persistence, so one should not expect autocorrelations close to one (Andrabi et al., 2011). Indeed, in Table 3 the Chinese and math test autocorrelations from 2000

---

[16] These correlations between noncognitive skills are somewhat overestimated because they contain common questions from the child questionnaire. Two questions for the 2000 self-esteem scale are used for the internalizing scale, and one question from the 2000 depression scale is used in the internalizing scale. One question is used in both internalizing and externalizing scales. Yet when the common questions are removed the correlations are still high: 0.816 becomes 0.777, 0.628 becomes 0.624 and 0.675 becomes 0.609.

[17] The cross-sectional correlations reported in Table 2 are well within the range of what has been reported in previous studies in the U.S. Using the 1988 National Educational Longitudinal Study (NELS) data, Eren and Ozbeklik (2013) reported a correlation coefficient of 0.23 between cognitive ability (10th grade math test score) and noncognitive ability (tenth grade Rosenberg Self-Esteem and Rotter Locus of Control Scales). Heckman et al. (2011) reported a correlation coefficient of 0.24 between cognitive factor and social-emotional factor, using 1979 National Longitudinal Survey of Youth (NLSY79) data. Using a data set that is representative of the Swedish population, Lindqvist and Vestman (2011) reported a correlation coefficient of 0.388 between cognitive ability and noncognitive ability.

to 2004 are only 0.094 and 0.110, respectively. In contrast, the 2000 cognitive ability test is more autocorrelated with later literacy test scores: 0.364 for the 2004 test and 0.290 for the 2009 test, and the literacy test autocorrelation from 2004 to 2009 is 0.393.

The five autocorrelations for noncognitive skills in Table 3 also are somewhat low, and only two are significantly different from zero at the 5% significance level. Specifically, the autocorrelation of the internalizing index from 2000 to 2004 is close to zero and statistically insignificant, and the same is true of the self-esteem scale from 2000 to 2009. The depressive symptoms scale is weakly autocorrelated from 2000 and 2009 (0.061, significant at the 10% level). Finally, the externalizing scale from 2000 and 2004 is more highly autocorrelated (0.114, highly significant).[18] In addition to measurement error, one possible reason for low autocorrelations in both noncognitive and cognitive skills is that in wave 2 (2004), the surveyed children were aged 13-16 and thus in the midst of adolescence, which can be a tumultuous time for their personalities. There may also be large variations in the timing and degree of physical and emotional maturation.

**B. Determinants of Education and Employment Status.** Before examining whether cognitive and noncognitive skills predict the school-to-work transition, Table 4 shows the types of jobs taken by those individuals who work. The vast majority (71.1%) are unskilled jobs. Among these, 20.3% are in manufacturing, which is more common for women (25.4%) than men (15.4%). Unskilled service sector jobs account for about a third of the jobs, and are divided into working in restaurants (12.8%), hotels, travel and entertainment (6.3%) and retail and other service sector jobs (14.0%). Unskilled service sector work is more common for women (about 44%) than for men (23%). More common for men, and rare for women, are jobs in construction and mining (9.0% of men), transportation (6.4%) and military service (5.5%). Of these young people who work, only 4.9% report being farmers, 3.5% of men and 6.4% of women.

The most common skilled worker category is "professionals", who are 27% of the sample of workers and do a wide range of jobs that require skills typically obtained in formal schooling or training courses. This includes teachers, nurses, computer technicians, machinery operators, mechanics, welders and bakers. Given the 17-21 year age range, it is very unlikely that many are teachers or nurses. Thus these jobs are most likely skilled production jobs, which accords with the fact that this type of work is more common for men (33.6%) than for women (20.1%).

Table 5 presents multinomial logit estimates of the school-to-work transitions observed in 2009, when the children were 17-21 years old. As explained above, the data include 1,863 of the original 2,000 children, of whom 847 (45.5%) are in school, 849 (45.6%) work, and 167 (9.0%) neither work nor are in school. In all regressions the base group is students.

The first two columns of Table 5 present estimates with only basic explanatory variables and three variables measuring cognitive skills (Chinese, math, and general cognitive skills) in the year 2000, when the children were 9-12 years old. As expected, students with higher levels of cognitive skills at age 9-12 are less likely to be working or idle (neither in school nor working)

---

[18] The autocorrelations reported in Table 3 are lower than in other studies. Using data collected in India, Helmers and Patnam (2011) reported that the correlation coefficient between the cognitive skill levels at age 8 and at age 12 is 0.9348. The correlation coefficient between the non-cognitive skill level at age 8 and the non-cognitive skill level at age 12 is 0.4276. A study at University of Michigan finds that the correlations between internalizing behavior reported by teachers when participants were 6 years old and 10 years old is 0.211 (Fan, 2011). The sample used is not representative of general population because it oversamples subjects with behavior problems. Using German adaptations of the Social Behavior Questionnaire (SBQ; Tremblay et al. 1987; Tremblay et al., 1992), Lösel and Stemmler (2012) find the correlation between girls' externalizing behavior when they were in elementary school and in secondary school is 0.2 and that between internalizing behavior is 0.32.

when they are 17-21 years old; thus they are more likely to be students. Turning to the other variables,[19] older individuals are more likely to be working or idle than being a student, girls are no more likely than boys to be in either category, and children with educated parents, especially educated fathers, are more likely to stay in school relative to working or being idle. Children from wealthier families also are much less likely to work than to be a student, but they are neither more nor less likely to be idle, relative to being a student.

The third and fourth columns of Table 5 address the main question of this study, whether - *noncognitive* skills predict students' school-to-work transitions over and above the predictive power of cognitive skills. Five such skills were measured in 2000, when the children were age 9-12: internalizing and externalizing behavior, educational aspirations, self-esteem and depressive symptoms.

The statistically significant negative coefficient on the internalizing scale suggests that relatively withdrawn students are more likely to stay in school. The externalizing scale is marginally significant positive for the outcome of working and is positive but insignificant for the outcome of being idle (both relative to being in school). Thus children who were more aggressive, destructive and/or impulsive when they are 9-12 years old are less likely to be in school at ages 17-21, even after accounting for their academic skills.

Students' educational aspirations are a strong predictor. As one would expect, students who desire to obtain a college or above education are less likely to be working. Somewhat surprisingly, greater self-esteem predicts a higher likelihood of working. Lastly, the depressive symptoms scale predicts higher probabilities of both employment and being idle at age 17-21. Overall, the evidence indicates that noncognitive skills have significant predictive power for the school-to-work transition over and above the predictive power of cognitive skills and of other variables.

The last two columns of Table 5 replace the cognitive and noncognitive skills measured in 2000 (when the children were 9-12 years old) with the same skills measured in 2004 (when they were 13-16). Of the three cognitive skill measures, the literacy test score, and to a lesser extent the math score, predict lower probabilities of both working and being idle relative to staying in school. Among the noncognitive skills, educational aspirations in 2000 and 2004 are very similar as predictors of staying in school. The only other noncognitive skill with a marginally statistically significant effect is resilience, which predicts a lower probability of working. The resilience indicator is composed of six subscales that measure optimism, emotional control, self-efficacy, interpersonal sensitivity, peer relationships, and relationships with adults. When the resilience scale is replaced by its six subscales (not shown in Table 5), only the estimated coefficient of optimism is statistically significant (5% level). Optimism predicts increased probability of staying in school. Overall, the 2004 data also show convincingly that noncognitive skills have predictive power independently of cognitive skills.

The bottom section of Table 5 reports the results of separate regressions that replace the individual skill measures with the cognitive and noncognitive factor variables. For brevity, only the results for the factor variables are reported; the estimated coefficients of the other variables change little. When the 2000 skills are used, both cognitive and noncognitive skills have predictive power, with better skills increasing the likelihood of staying in school rather than working or being idle. For the noncognitive factors, the 3$^{rd}$ factor (depression) has the strongest impact, while the 2$^{nd}$ factor (externalizing) does not have a significant impact. When the 2004

[19] Height-for-age Z-score, based on height measured in Wave 2, was also added as explanatory variable but it was never significant for labor force participation, nor for the wages.

skills are used one of the two cognitive factors and one of the two noncognitive factor influence the school-to-work transition. Overall, the results are in line with those that include specific skill measures.

Table 6 reports results from a linear probability model of whether children left school by 2009. Inspection of Table 5 reveals that the coefficient estimates are generally similar for working and being idle (relative to remaining in school), justifying our combining the two categories. The linear probability model has three advantages: one can add village fixed effects, one can employ instrumental variables in a way that makes identification transparent, and the coefficients can be interpreted as marginal probabilities. Linear probability models perform well (similar to non-linear models) when the mean probability is near 0.5, as in our case. Village fixed effects are needed because unobserved community characteristics may be correlated with the explanatory variables. For example, school quality may be correlated with children's cognitive skills when they are 9-12 years old and may also directly affect their education and employment decisions at age 17-21. Moreover, other community characteristics, such as local employment opportunities, may affect both noncognitive skills at an early age (uncertain parental earnings may affect children's psycho-social development) and later education and employment choices.

The first column of results in Table 6 is similar to the first set in Table 5, suggesting that grouping the working and idle is not very restrictive. In particular, after controlling for age, gender, parent education, and household wealth, students' cognitive skills (Chinese, math and general cognitive skills) in 2000 have strong predictive power: higher scores reduce the probability that a child is working or idle nine years later. The effects are large; a one standard deviation increase in the Chinese test score in 2000 predicts a 4.9 percentage points increase in the probability of being in school in 2009, and the effects are even larger for the 2000 math (6.0) and literacy (5.4) tests.

The results for noncognitive skills are also similar. More specifically, using measurements when the children were 9-12, college aspirations predict a decrease in the probability of working or being idle. Depressive symptoms and externalizing behavior predict increases in the same probability. The magnitudes of these effects are also large. Aspirations in 2000 to complete a college education increases the predicted probability of being in school in 2009 by 6.2 percentage points, and a one standard deviation increase in depressive symptoms increases the predicted probability by 4.9 percentage points . Two differences with the Table 5 results are that neither internalizing behavior nor self-esteem predicts the school-to-work transition; adding village fixed effects apparently reduced the predictive power of these variables.

The bottom of Table 6 show results of similar regressions that replace the multiple skill measures in 2000 with the cognitive and noncognitive factor variables. The estimated effects of the most significant cognitive and noncognitive factor are similar in magnitude. A one standard deviation increase in the 1st cognitive factor (cognition) predicts an increase in the probability of staying in school by 8.3 percentage points, while a one standard deviation increase in the 3rd noncognitive factor predicts an increase in the probability of staying in school by 9.5 percentage points. Estimated coefficients of both skill measures are statistically significant at the 1% level. Yet as noted earlier, because these skills are measured at an early age, and given the dynamic interactions between these two types of skills one cannot conclude that cognitive and noncognitive skills are equally important in determining labor force participation decisions at the time those decisions are made. We can only conclude that both sets of skills have predictive power at the time they are measured for future schooling and work outcomes.[20]

We also report results for a similar specification that replaces the 2000 cognitive and noncognitive skill variables with their 2004 counterparts (Table 6, column 4). As before, the Chinese, math and literacy test scores predict lower probabilities of leaving school. Interestingly, the effects of the 2004 Chinese and math tests are smaller than those for the 2000 Chinese and math tests, even though they were measured more recently. Perhaps test scores during adolescence are noisier measures of cognitive skills than measurements taken earlier in childhood. Also, the literacy test has a larger and more statistically significant coefficient than any of the other cognitive skill measures used in either year, suggesting that it better reflects the skills that matter for the school continuation decision. For these three test scores, a one standard deviation increase in the 2004 scores predicts an increase in the probability of being in school in 2009 by 2.9 to 8.0 percentage points. The size of the effects of the noncognitive skills measured in 2004 is again substantial. The aspiration to complete college or above education in 2004 predicts an increase in the probability of being in school in 2009 by 9.6 percentage points, and a one standard deviation increase in resilience (measured in 2004) predicts a 2.8 percentage points increase in that probability. Using the 2004 cognitive and noncognitive factor variables (bottom of Table 6), we find a strong and significant effect of the 2$^{nd}$ cognitive factor (literacy), much larger than for the 2000 cognitive factors, and a statistically significant but smaller impact of the 2$^{nd}$ noncognitive factor (resilience).

Columns 2-3 and 5-6 present gender specific regressions, which reveal some fascinating contrasts. First, the math and Chinese test scores are only strong predictors for girls; for boys the coefficients are much smaller and barely statistically significant only for math test scores. Thus, whether boys leave school is less related to academic performance than for girls. Yet boys' continuation in school is more strongly associated with higher cognitive ability (in 2000), relative to girls. The coefficient on the 2004 literacy test is also slightly more negative for boys than for girls. Turning to noncognitive skills, the most interesting finding is that 2004 resilience, measured during adolescence, predicts an increase in the probability of staying in school for boys, but not for girls. As in Table 5, the effect of resilience primarily reflects optimism. A one standard deviation increase in the optimism subscale predicts a decrease in the probability that

[20] In two alternative specifications of columns 1-3 in Table 6, child-reported measures of 2000 cognitive and noncognitive skills are replaced with mother-reported measures or teacher-reported measures. The results do reveal some differences in the predictive powers of alternative measures. The teacher-reported measure of cognitive ability has similar predictive power as the child-reported measure. The estimated coefficient of teacher-reported cognitive ability is somewhat larger in magnitude, but is not statistically different from the estimated coefficient of child-reported cognitive ability. The estimated coefficients of internalizing scale exhibit the largest differences. The estimated coefficient of mother-reported measure is an order of magnitude smaller than that of child-reported measure. In contrast, the estimated coefficient of teacher-reported measure is larger in magnitude than that of child-reported measure. In addition, it is also statistically significant in the male sample. Based on the teacher-reported measure, one standard deviation increase in the internalizing scale increases the likelihood of boys leaving school by 7.7 percentage points. The estimated coefficient of externalizing scale is large and more statistically significant using the child-reported measure than using mother-reported or teacher-reported measures. The teacher-reported measure of education aspiration is a stronger predictor than mother-reported or child-reported measures. The estimated coefficients of self-esteem are not statistically different across measures from different sources. The child-reported measure of depressed scale is a stronger predictor than mother-reported or teacher-reported measures. For 2004 skills, we do not have as many skills that have measures from alternative sources, especially for noncognitive skills. So we cannot conduct similar analysis as we have done for 2000 skills. Details of variable construction and regression results are available from the authors upon request.

boys leave school by 6.5 percentage points (subscale results not shown in Table 6). A one standard deviation rise in the optimism subscale also predicts a decrease in the likelihood that girls leave school by 3.9 percentage points. Also, a one standard deviation fall in the self-efficacy subscale predicts a decrease in boys' probability of leaving school by 4.1 percentage points (statistically significant at the 10% level). In contrast, the 2000 depressive scale is only an important predictor for girls but not for boys. Educational aspirations in 2000 or 2004 are strong predictors of staying in school for both girls and boys. Internalizing and externalizing scores are not statistically significant predictors in the gender-specific regressions. Overall, we find significant differences in the extent to which cognitive and noncognitive skills predict school-leaving behavior of girls versus boys, though as seen in the pooled regressions (columns 1 and 4 of Table 6) controlling for skills girls are no more likely to leave school than boys.[21]

As mentioned in Section III, it is very likely that both the cognitive and noncognitive skill variables have measurement error, so the results in Table 6 may underestimate those variables' effects on the school-to-work transition. There are several potential instruments for both types of skills to minimize attenuation bias. The wave 1 mother questionnaire asks each mother questions about her child's academic abilities and also asks exactly the same questions about her child that are used in the child questionnaire to construct the internalizing and externalizing scales. In addition, the child's homeroom teacher in 2000 was asked to complete a questionnaire on the child's academic skills and the child's personality and behavior.

Unfortunately, instrumenting for all the cognitive and noncognitive skill variables at once, or even just the cognitive skill variables at once, led to severe weak instruments problems (results not reported). Attempts to simultaneously instrument only the two latent skill variables led to the same problem. Thus we use the Andrews Moreira and Stock (2007) method; it adjusts standard errors to account for weak instruments but allows for only one endogenous variable.[22] We estimate specifications with only one skill measure at a time and use as IVs for each skill variable the questions asked of mothers and teachers that best reflect the skill measure being instrumented.[23] The results, shown in Appendix Table A.1, indicate much stronger predictive power when the cognitive or noncognitive skill variables are instrumented.[24] Thus the estimates in Table 6 very likely underestimate the true effects of both types of skills variables, perhaps by a large margin.

[21] When we use the 2000 cognitive and noncognitive factor variables, we find that for the first cognitive factor (cognition) and the first and third noncognitive factors (internalizing, depression), the magnitude of the effects are slightly greater for girls than for boys, but for the 2nd cognitive factor (math) the effects are larger for boys and not statistically significant for girls. Using the 2004 measures, we find that the 2nd cognitive factor (literacy) and the 2nd noncognitive factors (resilience) matter more for boys but that the 1st cognitive factor (Chinese/math) matters more for girls. Considering the score coefficients for the different factors, all of these results are easily understood, but do not appear to add greater power or insight compared to the specifications using specific skill measurements.

[22] This is implemented using the condivreg command in STATA written by Moreira and Poi (2001).

[23] Specific questions used as IVs for each skill measure are described in the notes to Appendix Table A.1. When estimating the impact of noncognitive skills (one at a time), the cognitive skills are included as controls because we are interested in the additional explanatory power of the noncognitive skills.

[24] Each coefficient reported in Table A.1 is from a separate regression that includes only one skill variable, plus the control variables used in Table 6. OLS estimates are shown for comparison. The *F*-statistics indicate the strength of the instruments, which are very weak for many of the noncognitive skill variables. Given that other skill variables are not controlled for in these regressions, one cannot infer much from the OLS and IV coefficients for any specific skill variable about the impacts of that specific skill; nonetheless the nature of the instruments and the consistency with which the OLS estimates are much smaller than the IV estimates provide supportive evidence that measurement error is causing downward bias in the OLS estimates in Table 6.

The estimates thus far provide evidence that "favorable" noncognitive skills (educational aspirations, self-esteem, absence of depressive symptoms, and resilience) when a child is 9-12 or 13-16 years old predicts whether he or she is in school at age 17-21, yet they do not indicate the extent to which higher noncognitive skills predict more years of schooling. One way to assess this is to estimate an ordered probit model of years of completed schooling that allows for censoring for those children who are still in school. Table 7 presents such estimates. We limit the sample to those in school in 2004 to preclude possible reverse causation if leaving school before 2004 influences skill measurements in 2004. Note that these estimates also likely underestimate true effects due to measurement error in the cognitive and noncognitive skill variables.

The first column in Table 7 shows that students with higher cognitive skills in 2000, when they were 9-12 years old and nearly all in school, have higher final years of schooling. This is as expected; students who do well in school stay longer, and China's upper secondary entrance exam very likely reinforces this pattern. More interesting for this paper is that noncognitive skills also play a role. Higher educational aspirations at a young age, and to some extent internalizing behavior, predict higher years of schooling attained. In contrast, higher self-esteem and greater depressive symptoms predict lower years of schooling. The self-esteem finding is surprising and its interpretation is unclear; separate regressions by gender indicate that the effect is primarily among girls. One possible explanation is that girls are more likely than boys to follow their parents' wishes that they stay in school, but this is less true for girls with higher self-esteem.

The second column in Table 7 replaces the 2000 cognitive and noncognitive skill variables with the 2004 variables. The results are similar, except that the Chinese achievement score loses its significance. Among noncognitive skill variables, education aspiration is a strong predictor. The resilience scale also predicts higher years of schooling obtained.

When we use the cognitive and noncognitive factors (bottom of Table 7), we find that both types of skills are strong predictors of years of schooling completed, and that this is the case for skills measured in both 2000 and 2004. For the 2004 measures, the most important cognitive factor—the 2$^{nd}$ cognitive factor (literacy)—has a much larger impact than the most important noncognitive factor—the 2$^{nd}$ noncognitive factor (resilience).

These effects of cognitive and noncognitive skills are large.[25] The estimates using the 2000 skills imply that one standard deviation increases in Chinese and math skills in 2000 predict increases in eventual years of schooling by 0.24 and 0.20 years, respectively. Similarly, having higher educational aspirations (desire to complete college or above education) predicts an increase in years of schooling by 0.19, while a one standard deviation increase in depressive symptoms predicts a decrease of -0.15 in years of schooling . For the 2004 skills, a one standard deviation increase in the literacy test predicts an increase in years of schooling by 0.30 years, and having higher education aspirations predicts an increase of 0.36 years.

**B. Determinants of Wages.** As explained in Section III, OLS estimates of earnings equations could suffer from selection bias, so a nonparametric sample selection correction method is employed to avoid this bias. Table A.2 presents three sets of probit estimates for whether an individual works for a wage. The results in the first column of Table A.2 show that the first identifying variable in the selection equation, failing the upper secondary school entrance exam, significantly increases the probability of working for a wage, as expected. The second identifying variable, bad harvests from 2000 to 2006, also increase that probability, significant at the 5% level. The second set of estimates in Table A.2 adds the 2000 test scores as explanatory variables.

[25] These calculations are available from the authors upon request.

The explanatory power of the two excluded variables is similar to that in the first column. Finally, the third column includes only observations with 2004 test scores; the results for the two excluded variables are similar to those in the first two columns. As explained in Section III, if the assumption that these two variables do not belong in the wage regression after conditioning on 2000 or 2004 skills does not hold, the estimated effects of cognitive and noncognitive skills on wages may still be biased.

Table 8 presents the basic wage regression results, as well as those for regressions that use cognitive and noncognitive skills measured in 2000 and 2004 as regressors. The first column includes no skill variables at all; it is a standard wage regression in which wages depend on years of schooling, years of experience, a dummy variable for women, basic household characteristics (schooling of parents and household wealth), village fixed effects, and a selection correction term (the selection propensity score from the first probit regression in Appendix Table A.2). The results are as one would expect. Young adults with an additional year of schooling receive wages that are 5.7% higher, and work experience is also associated with higher wages. Another finding that is not surprising (for China) is that women's wages are about 26% lower than those of comparable men. Three household background variables--father's education, mother's education and household per capita wealth in 2000--were also added to investigate whether children from better off households obtain higher paying jobs. Mother's education is associated with higher wages but the other two variables have small and statistically insignificant coefficients. Finally, the selection correction term is negative and statistically significant at the 5% level.

The remaining columns in Table 8 add cognitive and noncognitive skill variables to investigate whether noncognitive skills predict wage after controlling for cognitive skills (and years of education). Column 2 uses the skill variables that were collected in wave 1 (2000), while column 3 does the same for wave 2 (2004). In both cases, none of the cognitive and noncognitive skill variables has any predictive power, and the years of schooling variable changes little and maintains its statistical significance. None of the estimated parameters for cognitive and noncognitive skills is large, even allowing for some attenuation bias.[26] Regressions that use the cognitive and noncognitive factors instead of the specific skill measures (bottom of Table 8) also generally yield statistically insignificant results for the noncognitive factors. However, the 2000 1$^{st}$ cognitive factor (cognition) is marginally significant (at the 10% significance level), suggesting that a one standard deviation increase in this factor increases wages by 7%. In this case, reducing the number of skill variables and estimating latent factors may increase the power of the regression to find significant effects. However, the IV estimates (not shown in Table 8) that attempt to correct for measurement error in both types of skills do not yield any statistically significant coefficients of the cognitive and noncognitive skill variables.[27]

---

[26] In two alternative specifications of Column 2 in Table 8, child-reported measures of 2000 skills are replaced with mother-reported measures or teacher-reported measures. Neither yields statistically significant results. So child-reported, mother-reported or teacher-reported measures of 2000 skills do not seem to differ in their predictive powers of respondents' wages in 2009.

[27] When the sample is limited to individuals who are working the instruments are very weak; for the six cognitive skill variables only one has first-stage $F$-statistic greater than ten, and only two others have first-stage $F$-statistics greater than five, and for the nine noncognitive skill variables the highest $F$-statistic is 3.53. As a further check, we estimated instrumental variable specifications of the results shown in columns 2 and 3 that, instead of instrumenting all skill variables at once, instrumented each of them one at a time, as what's done in Appendix Table A.1, to see whether the statistically insignificant impacts in those two columns are due to trying to instrument many variables at once with rather weak instruments. In fact, even when each of the cognitive and noncognitive skill variables is the only endogenous variable in the regression, the estimated impacts are always statistically insignificant.

Finally, Tables A.5 and A.6 present OLS and IV wage regressions, respectively, that use cognitive and noncognitive skills that were measured contemporaneously with wages in 2009. Table A.3 presents the estimates that generate the selection correction terms for Tables A.5 and A.6. When years of schooling, work experience, and all three skills variables are included in the OLS model, the CES-D depression scale is statistically significant at the 10% level, and the other two skill variables are insignificant. Joint significance tests find that in the last two specifications the three skill variables are jointly significant at the 10% significance level and the two noncognitive skill variables are also jointly significant at the 10% level. When the noncognitive factor is used instead of the two noncognitive skill variables, we find that a one standard deviation increase in noncognitive skills increases wages by nearly 10%. Note that the years of schooling variable is significant in all of these regressions and the magnitude of its coefficient falls by less than 0.1 when all of the skill variables are included in the regression.

The estimated effects in Table A.5 could be biased due to measurement error and/or simultaneity bias. Unfortunately, attempts to instrument all three 2009 skill variables at the same time again led to a weak instruments problem for the noncognitive skills. As before, we adopt the method of Andrews, Moreira, and Stock (2006) to adjust the standard errors so that significance tests have the correct size in the presence of weak instruments. Once again we are limited to including one endogenous regressor at a time. All of the IV results thus should be interpreted cautiously.

Table A.6 presents the IV estimates, along with OLS estimates for the same specifications. Compared to the OLS results, the IV coefficient on the Rosenberg self-esteem scale is much more positive, and significant at the 10% level by itself but the statistical significance disappears when (non-instrumented) years of schooling and work experience are added. The IV estimates for the CES-D depression scale are more positive but statistically insignificant in contrast to the negative and statistically significant OLS coefficients. The noncognitive factor variable based on these two skills has slightly positive but statistically insignificant coefficients. The IV coefficient on the literacy (cognitive skill) has an unexpected negative sign, which is opposite of the OLS results. It also becomes statistically significant at the 10% level when years of schooling and work experience are added. It may reflect the fact that more skilled individuals enter occupations that initially pay lower wages but have better scope for wage increases in the future or otherwise are negatively selected into low-wage jobs. Adding occupation dummy variables (not reported) causes the coefficient to shrink from -0.35 to -0.26 and become statistically insignificant.

To summarize this subsection, Tables 8, A.5 and A.6 suggest that, other than one marginally significant effect of an aggregate cognitive factor in 2000, none of the cognitive skills measured earlier in life or concurrently with wages significantly predict wages after conditioning on years of schooling (and work experience). This does not necessarily imply that cognitive skills do not affect wages; it may simply indicate that in China years of schooling are closely correlated with cognitive skills and so adding potentially noisy measures of cognitive skills does not provide new information. In addition, the wage earners in this sample have just started working, and it is quite possible that cognitive skills will have more influence after employers learn more about the cognitive skills of their workers.

Turning to noncognitive skills, similar to the results of Heckman et al. (2011) for the US, there is no evidence in Table 8 that noncognitive skills measured when individuals are in primary school or during adolescence have predictive power for wages after conditioning on years of schooling and work experience. Yet this could simply reflect that those indicators of noncognitive skills have a large amount of noise; unfortunately the IV estimates are relatively imprecise and thus fail to provide strong evidence as to whether this could be the case. When contemporaneously measured noncognitive skills are used, there is some suggestive evidence that noncognitive skills matter, in particular self-esteem, but the effects are marginally significant and the estimation is done without controlling for other skill measures.

## V. Conclusion

Recent research using data from developed countries has shown that noncognitive skills may be as important as cognitive skills as predictors of the employment and wage outcomes of labor market participants. This paper is one of the first studies to investigate whether the same holds in developing countries, using evidence from rural China.

Our results indicate that noncognitive skills, measured at age 9-12, have additional predictive power beyond that provided by cognitive skills for the schooling-to-work transition at age 17-21. Children who have higher cognitive skills at age 9-12 are more likely to be in school at age 17-21, and ultimately complete more years of schooling. Cognitive skills are stronger predictors for girls than for boys. More interestingly, even after controlling for cognitive skills, children's noncognitive skills at age 9-12 and at age 13-16 predict their schooling-to-work transition at age 17-21. In particular children with higher educational aspirations and fewer depressive symptoms at age 9-12, and children who are more resilient at age 13-16, are more likely to be still in school at age 17-21 and ultimately complete more years of schooling. The depression scale is a strong predictor for girls, while the resilience effect is a stronger predictor for boys. A final result, which perhaps is counterintuitive, is that more self-esteem at age 9-12 predicts *fewer* years of completed schooling. This result is primarily found among girls; one possibility is that girls are more likely to follow their parents' advice to stay in school, but this is less so for girls with high self-esteem. Overall, we conclude that in poor, rural environments, noncognitive skills are an important factor affecting children's educational attainment.

The results of cognitive and noncognitive skills on wages are admittedly only suggestive. First, there is little evidence that cognitive skills, whether measured earlier in life or concurrently with wages, predict wages after conditioning on years of schooling. This does not necessarily imply that cognitive skills do not influence wages; it may simply indicate that in China years of schooling are closely correlated with cognitive skills and so adding possibly noisy measures of cognitive skills provides no new information. Second, there is no evidence that noncognitive skills measured when individuals are in primary school or during adolescence have predictive power for wages after conditioning on years of schooling, work experience and indicators of cognitive skills. When contemporaneously measured noncognitive skills are used, there is suggestive evidence that noncognitive skills, in particular self-esteem, matter.

Overall, this study provides a rich characterization of how cognitive and noncognitive skills measured during childhood and adolescence predicts educational attainment and early labor market outcomes in a poor, rural setting, demonstrating the value of collecting panel datasets in developing countries that measure multiple dimensions of cognitive and noncognitive skills and span school-to-work transitions. Different personality traits have different effects, some positive and some negative, and the effects vary significantly for boys versus girls.

One of the most significant findings of this study is that some noncognitive skills have strong predictive power for respondents' years of schooling. This is particularly important in the context of China. China's upper secondary attainment rate is less than one-third that of OECD countries and the lowest among the BRICS countries (Khor et al., 2015). Educational attainment is particularly low in poor, rural areas, as evidenced in our sample of rural youth. Among the respondents that have left school by 2009, 30% did not graduate from middle high school. About 41% graduated from middle high school but did not pursue any further education. Another 9% attended high school or equivalent school but did not graduate. In short, among the respondents that left school by 2009, 80% did not attend or graduate from high school. The low level of educational attainment is an issue China's leaders cannot afford to ignore since increasing the supply of skilled labor is crucial to support China's further economic growth and transition to an economy based on higher value-added, high-wage industries (Heckman and Yi, 2014; Khor et al., 2015).

Our study points to the usefulness of collecting measures of noncognitive skills during childhood to identify those with greater risk of low educational attainment, and if possible, alter those skills and future outcomes. Previous studies suggest that early interventions can affect noncognitive skills and the returns may be high. Simulation analysis of Cunha and Heckman (2008) show a larger share of increased earnings as the result of higher (parental) investment at various early ages (age 6-7, age 8-9, age 10-11) is attributed to higher noncognitive skills than higher cognitive skills. Thus, programs that target at boosting noncognitive skills or mitigating negative emotional factors could boost children's years of schooling. For example, other research using the GSCF data has found that parental migration adversely affects children's noncognitive skills (Lee and Park, 2011), suggesting the need for programs that provide greater support to such students.

Clearly, these results are only a first attempt to assess the role of cognitive and noncognitive skills on employment and wages in a developing county, and so there are several important areas for future research. First, more precise measures of both types of skills may provide stronger predictive power, as would a larger sample of wage earners. Second, the wage earners in this sample have just started working and exclude the more educated youth who are still in school, and so it is possible that skills will have more explanatory power for wages after workers gain more experience (and after employers learn more about the cognitive skills of their workers) and may matter more for more educated workers. Third, to the extent that noncognitive skills are important determinants of worker productivity, an important area of future research is the extent to which programs and policies can develop those skills in individuals, both when they are in school (and even before they enter school) and when they are adult.

**References**


Achenbach, T.M. and C.S. Edelbrock.1978. The Classification of Child Psychopathology: A Review and Analysis of Empirical Efforts. *Psychological Bulletin* 85: 1275-1301.

Almlund, M., A. Duckworth, J. Heckman, and T. Kautz. 2011. Personality Psychology and Economics, in E. Hanushek, S. Machin, and L. Woessman, eds., *Handbook of the Economics of Education*, Vol. 4, Amsterdam: Elsevier, pp. 1-181.

American Psychological Association. 2007. *APA Dictionary of Psychology*. Washington, D.C.: American Psychological Association.

Andrabi, Tahir, Jishnu Das, Asim Khwaja and Tristan Zajonc. 2011. "Do Value-Added Estimates Add Value? Accounting for Learning Dynamics". *American Economic Journal: Applied Economics* 3(3):29-54.

Andrews, Donald W.K., Marcelo J. Moreira, and James H. Stock. 2006. "Optimal Two-Sided Invariant Similar Tests for Instrumental Variables Regression," *Econometrica*. 74(3): 715-752.

Andrich, D. 1978. A rating formulation for ordered response categories. *Psychometrika* 43: 561-573.

Becker, G. 1964. *Human Capital*. New York: Columbia University Press (for National Bureau of Economic Research).

Boey K.W. 1999. Cross-validation of a short form of the CES-D in Chinese elderly. *Int J Geriatr Psychiatry* 14(8):608-17.

Borghans, L., A. Lee Duckworth, J. Heckman and B. ter Weel. 2008. "The Economics and Psychology of Personality Traits." *Journal of Human Resources* 43(4):972-1059.

Bowles, S., H. Gintis, and M. Osborne. 2001. "The Determinants of Earnings: A Behavioral Approach." *Journal of Economic Literature* 39(4): 1137–76.

Cai, H., J. D. Brown, C. Deng and M. A. Oakes. 2007. "Self-esteem and Culture: Differences in Cognitive Self-evaluations or Affective Self-regard?" *Asian Journal of Social Psychology,* 10, 162–170.

Card, D. 1999. "The Causal Effect of Education on Earnings," in O. Ashenfelter and D. Card, eds., *Handbook of Labor Economics Volume 3.* North Holland.

Card, D. 2001. "Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems," *Econometrica* 69(5):1127-1160.

Chen, X., L. Mowei, H. R. Kenneth. 2002. Parental Reports of Externalizing and Internalizing Behaviors in Chinese Children: Relevancy to Social, Emotional and School Adjustment, *Journal of Psychology in Chinese Societies*.

Chorpita, B. F., and D. H. Barlow. 1998. The development of anxiety: The role of control in the early environment. *Psychological Bulletin* 124: 3-21.

Cunha, F., and J. Heckman. 2008. Formulating, identifying, and estimating the technology of cognitive and noncognitive skill formation. *Journal of Human Resources* 43(4): 738–82.

Cunha, F., J. Heckman, and S. Schennach. 2010. Estimating the Technology of Cognitive and Noncognitive Skill Formation", *Econometrica* 78(3): 883-931.

Das, M., W. Newey, and F. Vella. 2003. "Nonparametric Estimation of Sample Selection Models". *Review of Economic Studies* 70: 33-58.

Das, J. and Zajonc, T. 2010. "India Shining and Bharat Drowning: Comparing two Indian States to the Worldwide Distribution in Mathematics Achievement", *Journal of Development Economics* 92(2): 175-187.

Dearing, E., K. McCartney and B. Taylor. 2006. Within-child Associations between Family Income and Externalizing and Internalizing Problems. *Developmental Psychology* 42(2): 237-252.

Eren, O.and S. Ozbeklik, 2013. "The effect of noncognitive ability on the earnings of young men: A distributional analysis with measurement error correction", *Labour Economics* 24: 293-304.

Evans, G. W. 2004. "The Environment of Childhood Poverty." *American Psychologist.* 59(2): 77-92.

Fan, 2011, C.K, 2011. A Longitudinal Examination of Children's Emotion Regulation Problems, Negative Parenting Behaviors, and the Development of Internalizing Behavior Problems, PhD dissertation, University of Michigan.

Glewwe, P. 2002. "Schools and Skills in Developing Countries: Education Policies and Socio-economic Outcomes." *Journal of Economic Literature* 40(2): 436-482.

Goldsmith, A. H., J. R. Veum, and W. Darity Jr. 1997. "The Impact of Psychological and Human Capital on Wages." *Economic Inquiry* 35(4): 815–29.

Hanushek, E. 2002. "Publicly Provided Education", in A. Auerbach and M. Feldstein, eds., *Handbook of Public Economics, Volume 4.* North Holland.

Hanushek, E, and L. Woessmann. 2008. "The Role of Cognitive Skills in Economic Development." *Journal of Economic Literature* 46(3): 607-668.

Heckman, J. 1979. "Sample Selection Bias as a Specification Error." *Econometrica* 47(1):153-160.

Heckman, J., J. E. Humphries, S. Urzua, and G. Veramendi. 2011. *The Effects of Educational Choices on Labor Market, Health, and Social Outcomes*, mimeo.

Heckman, J., J. Stixrud, and S. Urzua. 2006. "The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior." *Journal of Labor Economics* 24(3): 411–82.

Heckman, J. and J. Yi, 2014. "Human capital, economic growth, and inequality in China", in S. Fan, R. Kanbur, S. Wei, and X. Zhang, eds., *The Oxford Companion to the Economics of China*. Oxford University Press.

Helmers, C. and M. Patnam, 2011. "The formation and evolution of childhood skill acquisition: Evidence from India", *Journal of Development Economics* 95 (2): 252–266.

Ivanova, M. Y., T. M. Achenbach, L. A. Rescorla, et al. 2007a. The Generalizability of the Youth Self-Report Syndrome Structure in 23 Societies. *Journal of Consulting And Clinical Psychology* 75(5): 729-738.

Ivanova, M. Y., T. M. Achenbach, L. Dumenci, et al. 2007b. Testing the 8-syndrome structure of the child behavior checklist in 30 societies. Journal of Clinical Child And Adolescent Psychology 36(3): 405-417.

John, O. P., and S. Srivastava. 2009. "The Big Five Trait Taxonomy: History, Measurement and Theoretical Perspectives", in L.A. Pervin and O.P. John, Eds., Handbook of Personality: Theory and Research (New York: The Guilford Press, pp. 102-138.

Lee, L., and A. Park. 2011. Parental Migration and Child Development in Rural China, working paper, Oxford University.

Lindqvist, E. and R. Vestman. "The Labor Market Returns to Cognitive and Noncognitive Ability: Evidence from the Swedish Enlistment." *American Economic Journal: Applied Economics* 3: 101-128.

Liu, X. 2003. Parenting Practices and the Psychological Adjustment of Children in Rural China, Ph.D. dissertation, University of Pennsylvania

Lösel, F. and M. Stemmler. 2012. "Continuity and patterns of externalizing and internalizing behavior problems in girls: A valuable- and person oriented study from preschool to youth age", *Psychological Testing and Assessment Modeling* 54(3): 307-319.

Khor, N., L. Pang, C. Liu, F. Chang, D. Mo, P. Loyalka, and S. Rozelle, 2015. "China's Looming Human Capital Crisis: Upper secondary educational attainment rates and the middle Income Trap", Working Paper 280, Stanford University - Freeman Spogli Institute for International Studies.

Masters, G. 1982. A Rasch model for partial credit scoring. *Psychometrika* 47: 149-174.

Mincer, J. 1974. *Schooling, Experience and Earnings*. New York: Columbia University Press (for National Bureau of Economic Research).

Moreira, Marcelo J. and Brian P. Poi (2001). "Implementing Conditional Tests with Correct Size in the Simultaneous Equations Model," *Stata Journal*, Vol. 1, No. 1, pp. 1-15.

Noam G. G., and L. S. Goldstein. 1998. *The Resilience Inventory*, Unpublished Protocol.

OECD and Statistics Canada. 2000. "Literacy in the Information Age: Final Report of the International Adult Literacy Survey". Paris: OECD Publications.

Radloff, L.S. 1977. "The CES-D scale: A self report depression scale for research in the general population". *Applied Psychological Measurement* 1: 385-401.

Rasch, G. 1960. *Probabilistic Models for Some Intelligence and Attainment Tests*. Copenhagen: Danish Institute for Educational Research.

Rasch, G. 1961. On general laws and the meaning of measurement in psychology. In *Proceedings of the IV Berkeley Symposium on Mathematical Statistics and Probability*, ed. J. Neyman, 321-333. Berkeley, CA: University of California Press.

Rosenberg, M.. 1965. *Society and the Adolescent Self-Image*. Princeton, NJ: Princeton University Press.

Segal, C.. 2012 (forthcoming). "Misbehavior, Education, and Labor Market Outcomes", *Journal of the European Economic Association*.

Singh, Abhijeet. 2016. *Learning More With Every Year: School Year Productivity and International Learning Divergence*, working paper, University College London.

Song, H., H. Cai, J. D. Brown and K. J. Grimm. 2011. Differential item functioning of the Rosenberg Self-Esteem Scale in the US and China: Measurement bias matters. *Asian Journal of Social Psychology* 14(3): 176-188.

Song, M. 2001. "Psychological resilience among Korean adolescents: A measurement pilot on the Resilience Inventory." Unpublished qualifying paper, Harvard University, Cambridge.

Song, M. 2003. "Two Studies on the Resilience Inventory (RI): Toward the Goal ofreating a Culturally Sensitive Measure of Adolescence Resilience." Ph.D. Dissertation, School of Education, Harvard University.

Tremblay, R. E., Desmarais-Gervais, L., Ganon, C., & Charlebois, P. 1987. The Preschool Behavior Questionnaire. Stability of its factor structure between cultures, sexes, ages and socioeconomic classes. *International Journal of Behavioral Development, 10,* 467-484.

Tremblay, R. E., Vitaro, F., Gagnon, C., Piché, C., & Royer, N. 1992. A prosocial scale for the Preschool Social Behavior Questionnaire: Concurrent and predictive correlates. *International Journal of Behavioral Development, 15,* 227-245.

Vella, F. 1998. "Estimating Models with Sample Selection Bias: A Survey". *Journal of Human Resources* 33: 127-169.

Yen, S., C. J. Robins and N. Lin. 2000. A cross-cultural comparison of depressive symptom manifestation: China and the United States. *Journal of Consulting and Clinical Psychology* 68(6).

**Table 1. Cognitive and Noncognitive Skills in the Gansu Survey of Children and Families**

| *Year* | *Cognitive Skills* | *Noncognitive Skills* |
|---|---|---|
| 2000 (Wave 1) | 1. Chinese test [a]<br>2. Math test [a]<br>3. Cognitive Skills Test | 1. Internalizing behavior<br>2. Externalizing behavior<br>3. Educational aspiration (child aspires to complete college)<br>4. Self-esteem scale [b]<br>5. Depressive symptoms scale [c] |
| 2004 (Wave 2) | 1. Chinese test<br>2. Math test<br>3. Literacy test | 1. Internalizing behavior<br>2. Externalizing behavior<br>3. Educational aspiration<br>4. Resilience (with subscales in optimism, self-efficacy, adult relationship, peer relationship, interpersonal sensitivity and emotional control) |
| 2009 (Wave 3) | 1. Literacy test (similar to the one used in 2004) | 1. Rosenberg self-esteem<br>2. Depressive symptoms (from Center for Epidemiological Studies) |

a. Half of the 2000 sample were randomly selected to take the Chinese test and the other half to take the math test.

b. The self-esteem scale is constructed using the following questions from the 2000 survey: I have many things to be proud of; I always do things well; I always win praise from others for what I've done; I cannot do things well without my parents; I think I should be good at everything; I feel inferior to others; I am satisfied with my life; I have reasons for what I do. Reverse questions have been recoded (strongly agree changed to strongly disagree, agree changed to disagree).

c. The depression scale is constructed from the following questions from the 2000 survey: I am satisfied with my life; I have confidence in the future; I can live better than most people in the future; I won't feel very happy in my future life; I enjoy my life now; I often feel unhappy; I often feel lonely; I cannot concentrate on what I am doing; I think many people like me; I feel very happy. Reverse questions have been recoded (strongly agree changed to strongly disagree, agree changed to disagree).

## Table 2. Correlations between Cognitive and Noncognitive Skills

| **2000** | Chinese test score | Math test score | Cognitive ability test | Cognitive factor | Internalizing scale | Externalizing scale | Self-esteem scale | Depression scale | Educational aspiration |
|---|---|---|---|---|---|---|---|---|---|
| Chinese test score | 1 | | | | | | | | |
| Math test score | — [a] | 1 | | | | | | | |
| Cognitive ability test | 0.322*** | 0.247*** | 1 | | | | | | |
| Internalizing scale | -0.127*** | -0.0930*** | -0.248*** | -0.225*** | 1 | | | | |
| Externalizing scale | -0.138*** | -0.102*** | -0.273*** | -0.248*** | 0.816*** | 1 | | | |
| Self-esteem scale | -0.0198 | 0.0489* | -0.127*** | -0.0770*** | 0.114*** | 0.0265 | 1 | | |
| Depression scale | -0.147*** | -0.116*** | -0.335*** | -0.285*** | 0.675*** | 0.628*** | 0.146*** | 1 | |
| Educational aspiration | 0.114*** | 0.120*** | 0.214*** | 0.195*** | -0.128*** | -0.182*** | 0.0646** | -0.169*** | 1 |

a. Correlation between the Chinese and math test scores is not calculated for 2000 since children took either one or the other.

| **2004** | Chinese test score | Math test score | Literacy test | Cognitive factor | Internalizing scale | Externalizing scale | Resilience scale | Educational aspiration |
|---|---|---|---|---|---|---|---|---|
| Chinese test score | 1 | | | | | | | |
| Math test score | 0.485*** | 1 | | | | | | |
| Literacy test | 0.313*** | 0.310*** | 1 | | | | | |
| Internalizing scale | -0.0547* | -0.0492 | -0.0244 | -0.0573* | 1 | | | |
| Externalizing scale | -0.0897*** | -0.0825** | -0.114*** | -0.123*** | 0.745*** | 1 | | |
| Resilience scale | 0.153*** | 0.108*** | 0.180*** | 0.190*** | -0.387*** | -0.575*** | 1 | |
| Educational aspiration | 0.160*** | 0.152*** | 0.244*** | 0.238*** | -0.0883*** | -0.161*** | 0.215*** | 1 |

| **2009** | Literacy test | Rosenberg self-esteem scale | CES-D depression scale | Noncognitive factor |
|---|---|---|---|---|
| Literacy test | 1 | | | |
| Rosenberg self-esteem scale | 0.185*** | 1 | | |
| CES-D depression scale | -0.0595* | -0.262*** | 1 | |
| Years of schooling | 0.508*** | 0.194*** | -0.0742** | 0.169*** |

* denotes correlation is statistically significant at 10%, ** at 5% and *** at 1%.

**Table 3. Correlation over Time for Cognitive and Noncognitive Skills**

| | (2000, 2004) | (2004, 2009) | (2000, 2009) |
|---|---|---|---|
| Chinese test score | 0.094*** | | |
| Math test score | 0.110*** | | |
| Cognitive ability (2000)/ Literacy test score (2004 and 2009) | 0.364*** | 0.393*** | 0.290*** |
| Internalizing scale | 0.031 | | |
| Externalizing scale | 0.114*** | | |
| Self-esteem | | | -0.012 |
| Depression | | | 0.061* |
| Educational aspiration | 0.133*** | | |

* denotes correlation is statistically significant at 10%, ** at 5% and *** at 1%.

**Table 4. Types of Employment in 2009 for Individuals Who are Working (%)**

| *Type of Employment* | *Full Sample* | *Men* | *Women* |
|---|---|---|---|
| **Unskilled Jobs**: | 71.1 | 64.8 | 77.6 |
| Manufacturing | 20.3 | 15.4 | 25.4 |
| Restaurants | 12.8 | 10.1 | 15.6 |
| Hotel, travel, entertainment | 6.3 | 2.6 | 10.1 |
| Other jobs in service industry such as cashier, sales clerk, etc. | 14.0 | 9.9 | 18.3 |
| Construction or Mining [a] | 4.7 | 9.0 | 0.2 |
| Transportation | 3.5 | 6.4 | 0.5 |
| Own small business | 1.8 | 2.4 | 1.1 |
| Military | 2.8 | 5.5 | 0.0 |
| Farmer | 4.9 | 3.5 | 6.4 |
| **Skilled Jobs:** | 28.9 | 35.2 | 22.4 |
| Professionals [b] | 27.0 | 33.6 | 20.1 |
| Managers | 0.9 | 0.9 | 0.9 |
| Office jobs | 1.0 | 0.7 | 1.4 |

a. Most of the sample observations are in construction. Only a few are in mining.
b. This category is jobs that require skills (e.g., teachers, nurses, technicians, machine operators, welders, bakers).

**Table 5. School-to-Work Transitions using 2000 and 2004 Cognitive and Noncognitive Skills Multinomial Logit Model**

| | (1). 2000 Skills | | (2). 2000 Skills | | (3). 2004 Skills | |
|---|---|---|---|---|---|---|
| Base Group: Students | Working | Idle | Working | Idle | Working | Idle |
| Chinese achievement test score | -0.255*** | -0.294*** | -0.242** | -0.283*** | -0.100 | 0.028 |
| | (0.094) | (0.104) | (0.096) | (0.102) | (0.080) | (0.117) |
| Math achievement test score | -0.261*** | -0.249 | -0.251*** | -0.227 | -0.167** | -0.154 |
| | (0.068) | (0.161) | (0.070) | (0.158) | (0.077) | (0.126) |
| Cognitive ability test | -0.195** | -0.245** | -0.088 | -0.190 | | |
| | (0.078) | (0.117) | (0.076) | (0.121) | | |
| Literacy test score | | | | | -0.260*** | -0.456*** |
| | | | | | (0.082) | (0.122) |
| Internalizing scale (IRT) | | | -0.257** | -0.289* | -0.008 | -0.133 |
| | | | (0.128) | (0.174) | (0.087) | (0.164) |
| Externalizing scale (IRT) | | | 0.211* | 0.211 | -0.125 | -0.046 |
| | | | (0.116) | (0.158) | (0.106) | (0.165) |
| Educational aspiration | | | -0.396*** | -0.154 | -0.614*** | -0.786*** |
| | | | (0.118) | (0.169) | (0.158) | (0.208) |
| Self-esteem scale (IRT) 2000 | | | 0.143** | -0.0722 | | |
| | | | (0.059) | (0.101) | | |
| Depression scale (IRT) 2000 | | | 0.255*** | 0.230* | | |
| | | | (0.073) | (0.136) | | |
| Resilience scale (IRT) 2004 | | | | | -0.131* | -0.108 |
| | | | | | (0.070) | (0.116) |
| Age of sample children | 0.533*** | 0.342*** | 0.572*** | 0.368*** | 0.462*** | 0.273*** |
| | (0.045) | (0.069) | (0.047) | (0.071) | (0.051) | (0.071) |
| Gender dummy =1 if female | 0.124 | 0.041 | 0.120 | 0.033 | 0.030 | -0.130 |
| | (0.117) | (0.182) | (0.116) | (0.183) | (0.130) | (0.198) |
| Years of schooling of Father | -0.088*** | -0.079*** | -0.088*** | -0.078*** | -0.095*** | -0.058* |
| | (0.016) | (0.030) | (0.016) | (0.030) | (0.018) | (0.033) |
| Years of schooling of Mother | -0.045** | -0.032 | -0.044** | -0.034 | -0.030 | -0.011 |
| | (0.019) | (0.032) | (0.020) | (0.032) | (0.020) | (0.034) |
| Log of per capita wealth in 2000 | -0.249*** | -0.005 | -0.217*** | -0.001 | -0.261*** | -0.066 |
| | (0.069) | (0.110) | (0.068) | (0.109) | (0.074) | (0.127) |
| Constant | -7.573*** | -7.433*** | -8.359*** | -7.876*** | -5.693*** | -5.374*** |
| | (1.058) | (1.677) | (1.106) | (1.736) | (1.140) | (1.771) |
| 1st cognitive factor | -0.299*** | -0.354** | -0.195* | -0.290* | -0.178 | 0.136 |
| | (0.115) | (0.154) | (0.101) | (0.152) | (0.115) | (0.210) |
| 2nd cognitive factor | -0.218** | -0.277* | -0.199** | -0.265* | -1.358*** | -2.650*** |
| | (0.104) | (0.153) | (0.093) | (0.148) | (0.454) | (0.797) |
| 1st noncognitive factor | | | -0.159** | -0.081 | 0.093 | 0.149 |
| | | | (0.065) | (0.109) | (0.063) | (0.108) |
| 2nd noncognitive factor | | | 0.129 | -0.103 | -0.281*** | -0.417** |
| | | | (0.259) | (0.211) | (0.101) | (0.176) |
| 3rd noncognitive factor | | | -0.568*** | -0.465** | | |
| | | | (0.201) | (0.217) | | |

| | | | |
|---|---|---|---|
| Observations | 1862 | 1862 | 1561 |

a. Standard errors clustered at village level are reported in parentheses. * $p$<0.10, ** $p$<0.05, *** $p$<0.01.
b. Variables calculated from fitting Item Response Theory models are labeled with (IRT). The variable Resilience scale 2004 comes from fitting a confirmatory factor analysis model using the subscales of Resilience IRT variables. All IRT variables are standardized.
c. Chinese and math scores in 2000 and 2004, General cognitive ability 2000 and Literacy test score 2004 are standardized.
d. The coefficient on the dummy variable that equals one if a child took the Math test (as opposed to the Chinese test) is not reported.
e. Cognitive and Noncognitive factors are obtained using the exploratory factor analysis. Standard errors are bootstrapped to account for the fact that these variables are estimated.

**Table 6. School-to-Work Transitions using 2000 and 2004 Cognitive and Noncognitive Skills**
**Linear Probability Model with Village Fixed Effects**

| Dependent variable: | 2000 Skills | | | 2004 Skills | | |
|---|---|---|---|---|---|---|
| = 1 if Working or Idle; | (1) | (2) | (3) | (4) | (5) | (6) |
| = 0 if Student | All | Male | Female | All | Male | Female |
| Chinese achievement | -0.049*** | -0.005 | -0.092*** | -0.029** | -0.023 | -0.041** |
| test score | (0.018) | (0.024) | (0.023) | (0.014) | (0.022) | (0.017) |
| Math achievement | -0.060*** | -0.044* | -0.082*** | -0.029* | -0.020 | -0.053*** |
| test score | (0.016) | (0.025) | (0.025) | (0.016) | (0.024) | (0.017) |
| Cognitive ability (2000) | -0.054*** | -0.074*** | -0.021 | | | |
| | (0.018) | (0.025) | (0.025) | | | |
| Literacy test (2004) | | | | -0.080*** | -0.075*** | -0.066*** |
| | | | | (0.016) | (0.023) | (0.022) |
| Internalizing scale (IRT) | -0.031 | -0.030 | -0.034 | -0.014 | -0.030 | 0.0065 |
| | (0.024) | (0.031) | (0.030) | (0.019) | (0.029) | (0.026) |
| Externalizing scale (IRT) | 0.040* | 0.037 | 0.040 | -0.013 | -0.013 | -0.0069 |
| | (0.023) | (0.029) | (0.033) | (0.022) | (0.033) | (0.030) |
| Educational aspiration | -0.062*** | -0.053* | -0.066** | -0.096*** | -0.115*** | -0.079* |
| | (0.022) | (0.031) | (0.030) | (0.033) | (0.042) | (0.045) |
| Self-esteem scale (IRT) 2000 | 0.018 | 0.014 | 0.026 | | | |
| | (0.012) | (0.016) | (0.017) | | | |
| Depression scale (IRT) 2000 | 0.049*** | 0.037 | 0.057** | | | |
| | (0.016) | (0.022) | (0.023) | | | |
| Resilience scale (IRT) 2004 | | | | -0.028* | -0.043* | 0.005 |
| | | | | (0.015) | (0.022) | (0.023) |
| Age of sample children | 0.110*** | 0.097*** | 0.114*** | 0.089*** | 0.073*** | 0.098*** |
| | (0.008) | (0.013) | (0.010) | (0.009) | (0.015) | (0.011) |
| Gender dummy =1 if female | 0.020 | | | -0.003 | | |
| | (0.024) | | | (0.027) | | |
| Years of schooling of Father | -0.013*** | -0.014*** | -0.013*** | -0.014*** | -0.015*** | -0.014*** |
| | (0.003) | (0.004) | (0.004) | (0.003) | (0.004) | (0.004) |
| Years of schooling of Mother | -0.011*** | -0.013** | -0.009* | -0.007* | -0.011* | -0.006 |
| | (0.004) | (0.005) | (0.005) | (0.004) | (0.006) | (0.006) |
| Log of per capita | -0.026** | -0.031 | -0.031* | -0.026* | -0.021 | -0.030 |
| wealth in 2000 | (0.013) | (0.019) | (0.018) | (0.015) | (0.023) | (0.020) |
| Constant | -1.21*** | -0.908*** | -1.235*** | -0.791*** | -0.485 | -0.970*** |
| | (0.190) | (0.312) | (0.257) | (0.225) | (0.365) | (0.262) |
| Adjusted $R^2$ | 0.150 | 0.111 | 0.189 | 0.129 | 0.105 | 0.161 |
| 1st cognitive factor | -0.083*** | -0.083*** | -0.110*** | -0.022 | -0.005 | -0.076* |
| | (0.029) | (0.026) | (0.028) | (0.026) | (0.035) | (0.045) |
| 2nd cognitive factor | -0.034 | -0.080*** | -0.044 | -0.430*** | -0.414*** | -0.321** |
| | (0.027) | (0.031) | (0.027) | (0.108) | (0.144) | (0.145) |
| 1st noncognitive factor | -0.048*** | -0.037** | -0.045** | 0.020 | 0.029 | 0.006 |
| | (0.013) | (0.018) | (0.020) | (0.014) | (0.020) | (0.021) |
| 2nd noncognitive factor | 0.021 | 0.013 | 0.036 | -0.061*** | -0.092*** | -0.008 |
| | (0.044) | (0.045) | (0.056) | (0.021) | (0.030) | (0.031) |
| 3rd noncognitive factor | -0.095*** | -0.078 | -0.099* | | | |
| | (0.036) | (0.053) | (0.058) | | | |
| Adjusted $R^2$ | 0.143 | 0.112 | 0.182 | 0.127 | 0.105 | 0.159 |
| Observations | 1862 | 1005 | 857 | 1561 | 840 | 721 |

a. Standard errors clustered at village level are reported in parentheses. * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

b. Variables calculated from fitting Item Response Theory models are labeled with (IRT). The variable Resilience scale 2004 comes from fitting a confirmatory factor analysis model using the subscales of Resilience IRT variables. All IRT variables are standardized.

c. Chinese and math scores in 2000 and 2004, General cognitive ability 2000 and Literacy test score 2004 are standardized.

d. The coefficient on the dummy variable that equals one if a child took the Math test (as opposed to the Chinese test) is not reported.

e. Cognitive and Noncognitive factors are obtained using exploratory factor analysis. Standard errors are bootstrapped to account for the fact that these variables are estimated.

**Table 7. Years of Schooling, Censored Ordered Probit Model**
**Sample: Children Still in School in 2004**

| Dependent variable: = 1 if Highest degree ≤ Primary school<br>= 2 if start lower secondary, but did not finish<br>= 3 if finish lower secondary, and no further schooling<br>= 4 if entered upper secondary & years of schooling ≤10<br>= 5 if entered upper secondary & years of schooling =11<br>= 6 if entered upper secondary & years of schooling =12<br>= 7 if one or more years of post-secondary | (1)<br>2000<br>Skills | (2)<br>2004<br>Skills |
|---|---|---|
| Chinese achievement test score | 0.225*** | -0.006 |
| | (0.049) | (0.043) |
| Math achievement test score | 0.187*** | 0.100** |
| | (0.044) | (0.039) |
| Cognitive ability test | 0.102* | |
| | (0.052) | |
| Literacy test score | | 0.290*** |
| | | (0.045) |
| Internalizing scale (IRT) | 0.120* | 0.042 |
| | (0.067) | (0.052) |
| Externalizing scale (IRT) | -0.096 | 0.009 |
| | (0.062) | (0.055) |
| Educational aspiration | 0.173** | 0.354*** |
| | (0.069) | (0.084) |
| Self-esteem scale (IRT) 2000 | -0.069** | |
| | (0.029) | |
| Depression scale (IRT) 2000 | -<br>0.127*** | |
| | (0.043) | |
| Resilience scale (IRT) 2004 | | 0.067* |
| | | (0.035) |
| Age of sample children | -0.016 | -0.015 |
| | (0.027) | (0.025) |
| Gender dummy =1 if female | -0.108 | -0.049 |
| | (0.073) | (0.074) |
| Years of schooling of Father | 0.060*** | 0.057*** |
| | (0.010) | (0.011) |
| Years of schooling of Mother | 0.035*** | 0.033*** |
| | (0.010) | (0.010) |
| Log of per capita wealth in 2000 | 0.205*** | 0.228*** |
| | (0.038) | (0.043) |
| 1st cognitive factor | 0.307*** | -0.069 |
| | (0.049) | (0.088) |
| 2nd cognitive factor | 0.118* | 1.649*** |
| | (0.066) | (0.355) |
| 1st noncognitive factor | 0.067** | -0.031 |
| | (0.034) | (0.035) |
| 2nd noncognitive factor | -0.068 | 0.183*** |
| | (0.143) | (0.055) |
| 3rd noncognitive factor | 0.249** | |

| | | |
|---|---|---|
| | (0.119) | |
| Observations | 1588 | 1478 |

a. Standard errors clustered at village level are reported in parentheses. * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

b. Variables calculated using Item Response Theory models are labeled (IRT). The variable Resilience scale 2004 comes from fitting a confirmatory factor analysis model using the subscales of Resilience IRT variables. All IRT variables are standardized.

c. Chinese and math scores in 2000 and 2004, General cognitive ability 2000 and Literacy test score 2004 are standardized.

d. Coefficient on the dummy variable that equals one if a child took the Math test (as opposed to the Chinese test) is not reported.

e. Cognitive and Noncognitive factors are obtained using exploratory factor analysis. Standard errors are bootstrapped to account for the fact that these variables are estimated.

**Table 8. Wage Regressions using 2000 and 2004 Cognitive and Noncognitive Skills**
**OLS with Village Fixed Effects Model**

| | | 2000 Skills | 2004 Skills |
|---|---|---|---|
| Dependable variable: lnwage | (1) | (2) | (3) |
| Years of schooling | 0.057*** | 0.056*** | 0.068*** |
| | (0.015) | (0.015) | (0.021) |
| Work experience measured in years | 0.044** | 0.041** | 0.041* |
| | (0.017) | (0.018) | (0.022) |
| Chinese achievement test score | | 0.037 | 0.020 |
| | | (0.043) | (0.036) |
| Math achievement test score | | 0.013 | -0.018 |
| | | (0.042) | (0.034) |
| Cognitive ability test | | 0.030 | |
| | | (0.036) | |
| Literacy test score | | | -0.013 |
| | | | (0.037) |
| Internalizing scale (IRT) | | 0.009 | 0.051 |
| | | (0.050) | (0.040) |
| Externalizing scale (IRT) | | -0.025 | -0.038 |
| | | (0.045) | (0.046) |
| Educational aspiration | | -0.030 | 0.010 |
| | | (0.053) | (0.061) |
| Self-esteem scale (IRT) 2000 | | -0.013 | |
| | | (0.025) | |
| Depression scale (IRT) 2000 | | 0.026 | |
| | | (0.035) | |
| Resilience scale (IRT) 2004 | | | -0.022 |
| | | | (0.038) |
| Gender dummy =1 if female | -0.255*** | -0.260*** | -0.247*** |
| | (0.054) | (0.054) | (0.062) |
| Years of schooling of Father | 0.008 | 0.008 | 0.005 |
| | (0.008) | (0.008) | (0.011) |
| Years of schooling of Mother | 0.025*** | 0.025*** | 0.025** |
| | (0.009) | (0.009) | (0.011) |
| Log of per capita wealth in 2000 | -0.010 | -0.011 | -0.010 |
| | (0.028) | (0.029) | (0.036) |
| Pr(Observed wage) | 0.348** | 0.378** | 0.218 |
| | (0.173) | (0.168) | (0.196) |
| Constant | 0.637** | 0.636** | 0.617* |
| Adjusted $R^2$ | 0.089 | 0.089 | 0.071 |
| 1st cognitive factor | | 0.072* | 0.013 |
| | | (0.041) | (0.058) |
| 2nd cognitive factor | | -0.001 | -0.066 |
| | | (0.042) | (0.177) |
| 1st noncognitive factor | | -0.002 | -0.023 |
| | | (0.029) | (0.029) |
| 2nd noncognitive factor | | 0.016 | 0.019 |

| | | | |
|---|---|---|---|
| | | (0.048) | (0.047) |
| 3rd noncognitive factor | | -0.045 | |
| | | (0.065) | |
| Adjusted $R^2$ | | 0.088 | 0.073 |
| Observations | 813 | 813 | 654 |

a. All notes for Table 6 also apply to this table.

**Table A.1. School-to-Work Transitions, Specifications with One Endogenous Variable**

| | Dependent variable: = 1 if Working or Idle; = 0 if Student | OLS | IV | First stage *F*-statistic | Sample Size |
|---|---|---|---|---|---|
| 1 | Chinese achievement test score 2000 | -0.066*** | -0.355*** | 38.40 | 1785 |
| | | (0.018) | (0.0745) | | |
| 2 | Math achievement test score 2000 | -0.076*** | -0.474*** | 42.64 | 1785 |
| | | (0.017) | (0.0906) | | |
| 3 | General cognitive test score 2000 | -0.107*** | -0.479*** | 42.02 | 1785 |
| | | (0.018) | (0.0685) | | |
| 4 | Chinese achievement test score 2004 | -0.064*** | -0.342*** | 34.58 | 1575 |
| | | (0.013) | (0.0869) | | |
| 5 | Math achievement test score 2004 | -0.069*** | -0.711*** | 12.46 | 1575 |
| | | (0.014) | (0.246) | | |
| 6 | Literacy test score 2004 | -0.118*** | -0.516*** | 27.06 | 1564 |
| | | (0.016) | (0.0825) | | |
| 7 | Internalizing scale (IRT) 2000 | 0.034** | 1.161** | 2.11 | 1785 |
| | | (0.013) | (0.541) | | |
| 8 | Externalizing scale (IRT) 2000 | 0.046*** | 0.489*** | 5.97 | 1785 |
| | | (0.014) | (0.164) | | |
| 9 | Educational aspiration 2000 | -0.076*** | -1.200*** | 11.60 | 1783 |
| | | (0.022) | (0.464) | | |
| 10 | Self-esteem scale (IRT) 2000 | 0.017 | 1.597 | 0.13 | 1709 |
| | | (0.011) | (4.143) | | |
| 11 | Depressed scale (IRT) 2000 | 0.055*** | -0.311 | 0.29 | 1785 |
| | | (0.013) | (0.906) | | |
| 12 | Internalizing scale (IRT) 2004 | -0.011 | 1.545 | 0.36 | 1513 |
| | | (0.012) | (1.735) | | |
| 13 | Externalizing scale (IRT) 2004 | -0.005 | 0.794 | 1.46 | 1513 |
| | | (0.013) | (0.505) | | |
| 14 | Educational aspiration 2004 | -0.104*** | -0.732 | 6.62 | 1422 |
| | | (0.033) | (0.468) | | |
| 15 | Resilience scale (IRT) 2004 | -0.021 | -0.594 | 3.61 | 1513 |
| | | (0.014) | (0.383) | | |

a. Each line presents separate regressions. Each regression includes only one skill variable as the endogenous variable. Estimated coefficients on the control variables were similar to those in Table 6 and are not reported. Standard errors are adjusted for weak instrument using Andrews, Moreira and Stock (2007) method.

b. Instrument for Chinese achievement score 2000 (row 1) and 2004 (row 4) is Chinese score in Spring 2000 school exam; instrument for Math achievement score 2000 (row 2) and 2004 (row 5) is Math score in Spring 2000 school exam; instruments for General cognitive test score 2000 (row 3) and Literacy test score 2004 (row 6) are two dummy variables that equal one if teacher rated child's learning ability as excellent or as below average in 2000.

c. Instruments for 2000 (row 7) and 2004 (row 12) externalizing scale (IRT) are two internalizing scales, one based on 2000 mother assessment and the other based on 2000 teacher assessment; instruments for 2000 (row 8) and 2004 (row 13) externalizing scale (IRT) are two externalizing scales, one based on 2000 mother assessment and the other based on 2000 teacher assessment; Instrument for 2000 (row 9) and 2004 (row 14) educational aspiration is mother aspiration for child's highest education level in the same year; instrument for self-esteem scale (IRT) 2000 (row 10) is a dummy variable that equals one if mother agrees with statement that child cannot do things well without the presence of parents; instrument for Depressed scale (IRT) 2000 (row 11) is a dummy variable that equals one if teacher in 2000 says child is sad or depressed; instrument for resilience (IRT) 2004 (row 15) is a dummy variable that equals one if teacher in 2000 says child appears to have given up in school.

**Table A.2. School-to-Work Transitions**
**Probit Models that Generate Selection Correction Terms**

| Dependent variable =1 if observed wage | (1) | (2) 2000 Skills | (3) 2004 Skills |
|---|---|---|---|
| Failed the entrance exam to high school | 0.862*** | 0.818*** | 0.884*** |
| | (0.0921) | (0.0949) | (0.100) |
| Years of bad harvest during 2000-2006 | 0.0572** | 0.0557** | 0.0638** |
| | (0.0230) | (0.0232) | (0.0259) |
| Chinese achievement test score | | -0.106* | -0.0757 |
| | | (0.0624) | (0.0491) |
| Math achievement test score | | -0.123** | -0.0685 |
| | | (0.0548) | (0.0544) |
| Cognitive ability (2000)/Literacy test score(2004) | | -0.0700 | -0.132*** |
| | | (0.0553) | (0.0506) |
| Internalizing scale (IRT) | | -0.0929 | 0.0139 |
| | | (0.0724) | (0.0630) |
| Externalizing scale (IRT) | | 0.0933 | -0.0297 |
| | | (0.0653) | (0.0688) |
| Educational aspiration | | -0.168*** | -0.165 |
| | | (0.0620) | (0.105) |
| Self-esteem scale (IRT) 2000 | | 0.0734* | |
| | | (0.0377) | |
| Depressed scale (IRT) 2000 | | 0.0923* | |
| | | (0.0546) | |
| Resilience scale (IRT) 2004 | | | -0.0487 |
| | | | (0.0464) |
| Age of sample children | 0.268*** | 0.310*** | 0.275*** |
| | (0.0277) | (0.0304) | (0.0335) |
| Gender dummy =1 if female | 0.0510 | 0.0383 | -0.00826 |
| | (0.0764) | (0.0752) | (0.0857) |
| Years of schooling of Father | -0.0376*** | -0.0346*** | -0.0447*** |
| | (0.0104) | (0.0104) | (0.0125) |
| Years of schooling of Mother | -0.0377*** | -0.0337*** | -0.0265* |
| | (0.0129) | (0.0129) | (0.0138) |
| Log of per capita wealth in 2000 | -0.0647 | -0.0379 | -0.0530 |
| | (0.0487) | (0.0483) | (0.0551) |
| Constant | -4.348*** | -5.117*** | -4.405*** |
| | (0.662) | (0.694) | (0.797) |
| 1st cognitive factor | | -0.132** | -0.098 |
| | | (0.057) | (0.085) |
| 2nd cognitive factor | | -0.074 | -0.674** |
| | | (0.066) | (0.309) |
| 1st noncognitive factor | | -0.076* | -0.001 |
| | | (0.044) | (0.045) |

| | | | |
|---|---|---|---|
| 2nd noncognitive factor | | 0.063 | -0.081 |
| | | (0.065) | (0.067) |
| 3rd noncognitive factor | | -0.212*** | |
| | | (0.068) | |
| Observations | 1851 | 1850 | 1534 |

All notes for Table 6 also apply to this table.

**Table A.3. Whether Wages and 2009 Skills are Observed**
**Bivariate Probit Model that Generates Selection Correction Terms**

| | (1)<br>=1 if observed wage | (2)<br>=1 if observed all 2009 skills |
|---|---|---|
| Failed the entrance exam to high school | 0.852*** | -0.330*** |
| | (0.0905) | (0.0801) |
| Years of bad harvest during 2000-2006 | 0.0549** | -0.0167 |
| | (0.0230) | (0.0305) |
| Age of sample children | 0.265*** | -0.144*** |
| | (0.0277) | (0.0265) |
| Gender dummy =1 if female | 0.0493 | -0.0713 |
| | (0.0766) | (0.0808) |
| Years of schooling of Father | -0.0361*** | -0.00175 |
| | (0.0101) | (0.0113) |
| Years of schooling of Mother | -0.0367*** | 0.0510*** |
| | (0.0129) | (0.0123) |
| Log of per capita wealth in 2000 | -0.0666 | 0.0436 |
| | (0.0485) | (0.0476) |
| Constant | -4.309*** | 3.057*** |
| | (0.667) | (0.637) |
| Correlation (ρ) | -0.419*** | |
| | (0.0528) | |
| Observations | 1863 | |

**Table A.4. First Stage Regressions of the IV Wage Regressions in Table A.6.**

| | (1) Rosenberg Self-esteem scale (IRT) | (2) CES-D Depression scale (IRT) | (3) Literacy test score | (4) Factor Noncognitive |
|---|---|---|---|---|
| Chinese achievement test score in 2000 | -0.051 | -0.173* | | |
| | (0.092) | (0.096) | | |
| Math achievement test score in 2000 | 0.140 | -0.081 | | |
| | (0.105) | (0.089) | | |
| Took Math achievement test 2000 | -0.145 | -0.046 | | |
| | (0.096) | (0.106) | | |
| General cognitive test score 2000 | 0.166** | 0.022 | | |
| | (0.083) | (0.080) | | |
| Internalizing scale (IRT) 2000 | 0.023 | -0.093 | | |
| | (0.109) | (0.125) | | |
| Externalizing scale (IRT) 2000 | -0.108 | 0.106 | | |
| | (0.088) | (0.098) | | |
| Educational aspiration 2000 | -0.066 | 0.058 | | |
| | (0.130) | (0.135) | | |
| Self-Esteem scale (IRT) 2000 | 0.021 | -0.020 | | |
| | (0.061) | (0.060) | | |
| Depressed scale (IRT) 2000 | 0.137 | 0.102 | | |
| | (0.084) | (0.083) | | |
| Literacy test score 2004 | | | 0.210*** | |
| | | | (0.064) | |
| 2000 1st cognitive factor | | | | 0.118 |
| | | | | (0.079) |
| 2000 1st cognitive factor | | | | 0.110 |
| | | | | (0.077) |
| 2000 1st noncognitive factor | | | | 0.021 |
| | | | | (0.049) |
| 2000 2nd noncognitive factor | | | | 0.062 |
| | | | | (0.073) |
| 2000 3rd noncognitive factor | | | | -0.030 |
| | | | | (0.094) |
| 2004 1st cognitive factor | | | | -0.018 |
| | | | | (0.081) |
| 2004 1st cognitive factor | | | | -0.460 |
| | | | | (0.292) |
| 2004 1st noncognitive factor | | | | 0.030 |
| | | | | (0.038) |
| 2004 2nd noncognitive factor | | | | 0.119* |
| | | | | (0.071) |
| Years of schooling | 0.083** | -0.011 | 0.051 | 0.051* |

| | | | | |
|---|---|---|---|---|
| | (0.036) | (0.043) | (0.037) | (0.028) |
| Work experience measured in years | -0.011 | 0.087** | -0.026 | -0.025 |
| | (0.034) | (0.039) | (0.036) | (0.022) |
| Gender dummy =1 if female | 0.089 | -0.054 | 0.146 | 0.027 |
| | (0.111) | (0.135) | (0.092) | (0.076) |
| Years of schooling of Father | -0.028** | 0.007 | -0.004 | -0.018* |
| | (0.014) | (0.014) | (0.014) | (0.009) |
| Years of schooling of Mother | -0.008 | 0.006 | -0.002 | -0.003 |
| | (0.021) | (0.023) | (0.017) | (0.015) |
| Log of per capita wealth in 2000 | 0.119* | -0.117 | -0.052 | 0.092* |
| | (0.066) | (0.077) | (0.060) | (0.048) |
| Pr(Observed wage but not 2009 skills) | -0.828 | 1.952* | -0.075 | -1.037 |
| | (1.048) | (1.110) | (1.103) | (0.718) |
| Pr(Observed 2009 skills but not wage) | -0.439 | 1.359* | 0.014 | -0.580 |
| | (0.616) | (0.733) | (0.688) | (0.421) |
| Observations | 395 | 395 | 395 | 395 |

a. All notes to Table 8 also apply to this table.

**Table A.5. Wage Regressions using 2009 Cognitive and Noncognitive Skills OLS with Village Fixed Effects Model**

| Dependent variable: lnwage | (1) | (2) |
|---|---|---|
| Years of schooling | 0.067** | 0.058* |
| | (0.031) | (0.032) |
| Work experience measured in years | 0.038 | 0.043 |
| | (0.031) | (0.031) |
| Rosenberg Self-esteem scale (IRT) 2009 | | 0.027 |
| | | (0.048) |
| CES-D Depression scale (IRT) 2009 | | -0.084* |
| | | (0.050) |
| Literacy test score 2009 | | 0.028 |
| | | (0.048) |
| Gender dummy =1 if female | -0.389*** | -0.397*** |
| | (0.078) | (0.078) |
| Years of schooling of Father | 0.007 | 0.008 |
| | (0.020) | (0.020) |
| Years of schooling of Mother | 0.031** | 0.032** |
| | (0.015) | (0.015) |
| Log of per capita wealth in 2000 | 0.004 | -0.008 |
| | (0.047) | (0.049) |
| Pr(Observed wage but not 2009 skills) | -0.519 | -0.325 |
| | (0.779) | (0.777) |
| Pr(Observed 2009 skills but not wage) | -0.268 | -0.133 |
| | (0.465) | (0.456) |
| Constant | 0.929 | 1.001* |
| | (0.594) | (0.593) |
| Adjusted $R^2$ | 0.122 | 0.139 |
| 2009 noncognitive factor | | 0.133** |
| | | (0.066) |
| Adjusted $R^2$ | | 0.139 |
| Observations | 402 | 402 |

a. Standard errors clustered at the village level are reported in parentheses. * $p$<0.10, ** $p$<0.05, *** $p$<0.01.
b. Rosenberg Self-esteem and CES-D Depression 2009 scales are calculated from fitting Item Response Theory models.
c. Rosenberg Self-esteem, CES-D Depression scale and Literacy test score 2009 are standardized.
d. Noncognitive factor is obtained using exploratory factor analysis. Standard error is bootstrapped to account for the fact that it is an estimated variable.

**Table A.6. OLS and IV Wage Regressions with 2009 Skills (village fixed effects)**

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Dependent variable: lnwage | OLS | OLS | IV | IV |
| Rosenberg Self-esteem scale (IRT) 2009 | 0.072* | 0.058 | 0.310* | 0.304 |
| | (0.040) | (0.040) | (0.179) | (0.248) |
| First stage *F*-statistic | | | 2.35 | 0.80 |
| CES-D Depression scale (IRT) 2009 | -0.098** | -0.093** | -0.010 | 0.144 |
| | (0.041) | (0.042) | (0.137) | (0.178) |
| First stage *F*-statistic | | | 1.73 | 1.19 |
| Literacy test score 2009 | 0.047 | 0.039 | -0.106 | -0.336* |
| | (0.042) | (0.040) | (0.148) | (0.194) |
| First stage *F*-statistic | | | 16.96 | 18.56 |
| Noncognitive factor 2009 | 0.155** | 0.128* | 0.384 | 0.364 |
| | (0.070) | (0.071) | (0.251) | (0.268) |
| First stage *F*-statistic | | | 2.53 | 1.15 |
| Control for years of schooling and experience? | No | Yes | No | Yes |
| Observations | 405 | 402 | 405 | 402 |

a. Standard errors clustered at the village level and in parentheses. * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

b. Rosenberg Self-esteem and CES-D Depression scale are calculated from fitting Item Response Theory models. Noncognitive factor 2009 is obtained using exploratory factor analysis. Standard errors for the OLS regressions are bootstrapped to account for the fact that it is an estimated variable.

c. Rosenberg Self-esteem scale 2009, CES-D Depression scale 2009, Literacy test score 2009 and Noncognitive factor 2009 are standardized.

d. Instruments for Rosenberg Self-esteem scale 2009 are 2000 Cognitive and Noncognitive skills. Instruments for CES-D Depression scale 2009 are 2000 Cognitive and noncognitive skills. Instrument for Literacy test score 2009 is literacy test score 2004. Instruments for Noncognitive factor 2009 are factors of the 2000 and 2004 Cognitive and Noncognitive skills.

e. All regressions control for gender, father's years of schooling, mother's years of schooling, log of per capita wealth in 2000, inverse mills ratio for wage, inverse mills ratio for 2009 skills, and village fixed effects (all variables differenced from village means).